SAND REPORT

SAND2004-1173
Unlimited Release
Printed March 2004

Obstacle Detection for Autonomous Navigation: An LDRD Final Report

Denise D. Padilla

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/ordering.htm

SAND2004-1173
Unlimited Release
Printed March 2004

Obstacle Detection for Autonomous Navigation: An LDRD Final Report

Denise D. Padilla
Intelligent Systems and Robotics Center

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1003

Abstract

This report summarizes the analytical and experimental efforts for the Laboratory Directed Research and Development (LDRD) project entitled "Obstacle Detection for Autonomous Navigation". The principal goal of this project was to develop a mathematical framework for obstacle detection. The framework provides a basis for solutions to many complex obstacle detection problems critical to successful autonomous navigation. Another goal of this project was to characterize sensing requirements in terms of physical characteristics of obstacles, vehicles, and terrain. For example, a specific vehicle traveling at a specific velocity over a specific terrain requires a sensor with a certain range of detection, resolution, field-of-view, and sufficient sensitivity to specific obstacle characteristics. In some cases, combinations of sensors were required to distinguish between different hazardous obstacles and benign terrain. In our framework, the problem was posed as a multidimensional, multiple-hypothesis, pattern recognition problem. Features were extracted from selected sensors that allow hazardous obstacles to be distinguished from benign terrain and other types of obstacles. Another unique thrust of this project was to characterize different terrain classes with respect to both positive (e.g., rocks, trees, fences) and negative (e.g., holes, ditches, drop-offs) obstacles. The density of various hazards per square kilometer was statistically quantified for different terrain categories (e.g., high desert, ponderosa forest, and prairie). This quantification reflects the scale, or size, and mobility of different types of vehicles. The tradeoffs between obstacle detection, position location, path planning, and vehicle mobility capabilities were also to be characterized.

Acknowledgements

This author wishes to thank the following individuals for their invaluable contributions, time, extensive discussions on the project, and input to this final report:

Jeff Carlson, Department 15211

Maritza Muguira, Department 15211

Jason Neely, Department 15211

Barry Spletzer, Department 15200

Dave Novick, Department 15211

Joel Beer, Department 15211

Patrick Davidson, Department 15211

Joshua Fowler, Department 15211

Jon Bryan, Department 2331

Steve Eskridge, currently at the Commercial Space Center for Engineering at Texas A&M

Table of Contents

Table of Figures

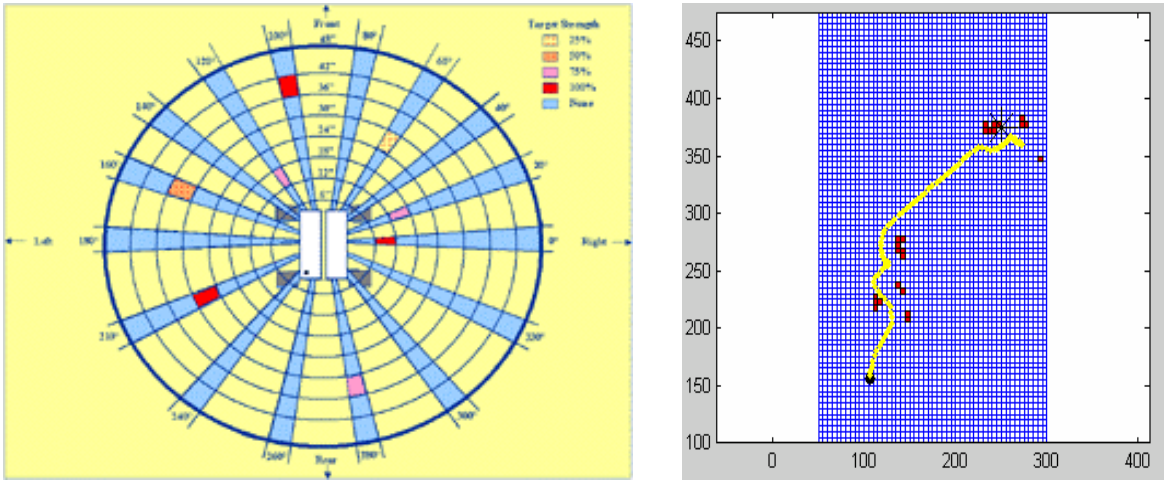
Table of Figures

Introduction

Since the terrorists acts of 9/11 there is an ever-increasing need for autonomously navigating vehicles for counter-terrorism and defense missions. The limiting factor in autonomous navigation is the ability to sense and map the environment. The principal goal of this project was to develop a mathematical framework for obstacle detection. The framework provides a basis for solutions to many complex obstacle detection problems critical to successful autonomous navigation. Another goal of this project was to characterize sensing requirements in terms of physical characteristics of obstacles, vehicles, and terrain. For example, a specific vehicle traveling at a specific velocity over a specific terrain requires a sensor with a certain range of detection, resolution, field-of-view, and sufficient sensitivity to specific obstacle characteristics. In some cases, combinations of sensors were required to distinguish between different hazardous obstacles and benign terrain. In our framework, the problem was posed as a multidimensional, multiple-hypothesis, pattern recognition problem. Features were extracted from selected sensors that allow hazardous obstacles to be distinguished from benign terrain and other types of obstacles. Another unique thrust of this project was to characterize different terrain classes with respect to both positive (e.g., rocks, trees, fences) and negative (e.g., holes, ditches, drop-offs) obstacles. The density of various hazards per square kilometer was statistically quantified for different terrain categories (e.g., high desert, ponderosa forest, and prairie). This quantification reflects the scale, or size, and mobility of different types of vehicles. The tradeoffs between obstacle detection, position location, path planning, and vehicle mobility capabilities were also to be characterized. Another major effort involved the investigation and evaluation of different sensor technologies for obstacle detection. We investigated several technologies including an acoustic imager, a modified ultrasonic system, structured lighting, scanning laser rangefinders, stereo vision, and LADAR.

Obstacle Detection Framework

We proposed to jump-start a new approach to autonomous navigation research by developing a systematic framework for obstacle detection. Principles of classical decision and estimation theory provide a mathematical basis for this framework. A Weighted-Nearest-Neighbor (WNN) decision model is a major feature. The WNN model is used for discovering and fusing features to achieve near-minimum probability of error decision algorithms. This framework provides the basis for solutions to a broad class of complex obstacle detection problems critical to successful autonomous navigation. Part of this framework includes an architecture for both sensor and control algorithm development. The architecture defines standard data structures including a local, vehicle-based, occupancy grid and a global, world-based, occupancy grid. These data structures provide sensor systems engineers specifications to design to and control and path planning systems engineers specifications to design from. An ultrasonic obstacle detection system, using commercially available sensors, was implemented using the architectural framework and was ported to a mobile robotic platform. The vehicle based occupancy grid is shown below in Figure 1a. Note the grid is based in polar coordinates and is relatively sparse -- composed of instantaneous sensor measurements. The grid is updated on the order of ten times per second. It shows the positions of obstacles relative to the vehicle and also indicates their density. A dense world-based grid shown in Figure 1b is constructed by transforming a sequence of instantaneous vehicle-based measurements using vehicle kinematics information (e.g., position and orientation). The world-based grid is used for vehicle control and path planning.



(a) (b)

Figure 1. Vehicle Based and World Based Occupancy Grids

Terrain and Obstacle Characterization

In our approach, we developed the *a priori* knowledge, or knowledge base, critical to the obstacle detection problem. A major part of our work this year involved the characterization of obstacles, terrain, and sensing requirements. In completing the obstacle characterization milestone, several classes of obstacles were characterized including solids such as rocks and trees, non-solids such as bushes and weeds, and negative obstacles, i.e., holes or drop-offs in the path of the mobile robot. This task defined the critical physical attributes of obstacles such as size, texture, reflectivity and the density with which they appear in the real-world workspace of the mobile robot. We initially surveyed the obstacles found in different areas surrounding the Robot Vehicle Range (RVR) that represent

high desert terrain. The picture below shows some of the typical obstacles such as tumbleweeds, bushes, and link fences found in this difficult to navigate terrain.

Figure 2. Typical obstacles seen at the RVR

In completing our terrain characterization milestone, various areas (grassy and hilly) were sampled and analyzed. A 300-foot tape measure was stretched across each sample area and, the height and length of the foliage or ground and other obstacles directly underneath the tape were documented. With this data, we were able to extract statistics characterizing different terrain. For example, the histogram below shows the density, or percentage of occurrence of obstacles versus their height for the specified terrain. This data is critical in the design of a vehicle in order to determine the probability of success for given terrain and mobility capabilities.

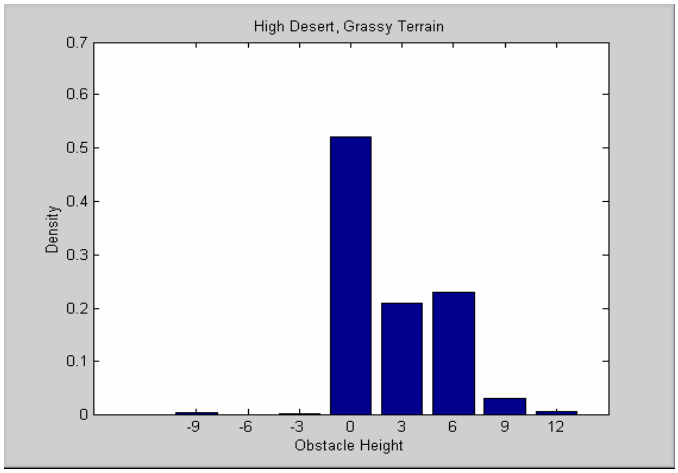


Figure 3. Histogram illustrating the density of obstacles versus their height

Evaluation of Sensor Technologies for Obstacle Detection

Another major effort involved the investigation and evaluation of different sensor technologies for obstacle detection. We investigated several technologies including an acoustic imager, a modified ultrasonic system, structured lighting, scanning laser rangefinders, stereo vision, and LADAR.

Modified Ultrasonic System

One of our sensor efforts involved the investigation of a modified ultrasonic sensor. Conventional ultrasonic ranging systems simply output a range measurement based on the time of flight of the first return to exceed a set threshold. These systems are intended for operation in controlled environments and are commonly used to detect the presence or absence of objects. The detection threshold is set to detect even the slightest return. In autonomous navigation, a blade of grass must be discriminated from a boulder or wall. Therefore, we looked at the entire echo waveform to classify different types of obstacles. The plots below in Figure 4 show a marked difference between solid objects such as a tire and penetrable objects such as grass and bushes. Note that the first response you see is the transmitted acoustic pulse. The half-power beam width of this particular sensor is approximately 8 degrees. The return over the grassy area is spread out over time because portions of the beam penetrate through to more distant features in the terrain. Solid obstacles such as the tire yield a concentrated return over a much narrower time span. We used these characteristics to develop a modified sensor with detection algorithms specifically optimized for obstacle detection purposes important to autonomous navigation.

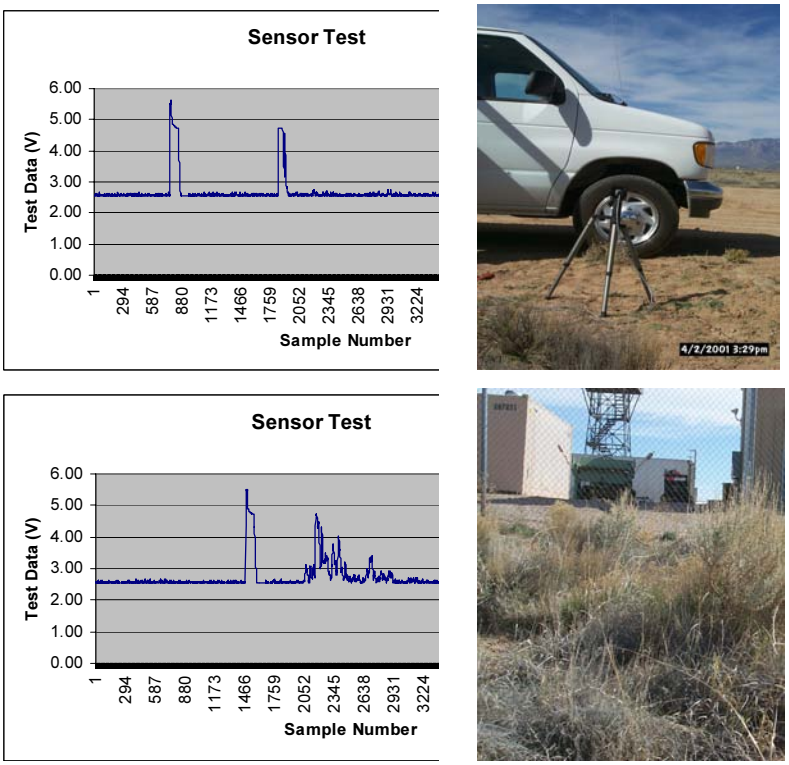


Figure 4. Ultrasonic echo returns

We collected ultrasonic echo returns from negotiable areas such as clear ground and thick grassy areas and nonnegotiable obstacles such as rocks, trees, fences, and posts. Intuitively, we searched for peaks in the echo return that maintain a high pulse for at least two samples. The ultrasonic sensor provides an amplitude vector for the echo response with values ranging from 0 to 5 volts. In order to identify the optimum decision rule, we found the maximum likelihood estimates for the amplitude vector of the echo response for negotiable paths and for obstacles. We analyzed the statistics gathered from our sample collection to determine the optimum features for the dichotomy of negotiable path and obstacles. The likelihood ratio is seen below.

$$\Lambda(\bar{x}) = \frac{f(\bar{x}|H_o)}{f(\bar{x}|H_a)}$$

If $\Lambda(\bar{x}) > \eta$, then the entire sensor return came from background (i.e. clear ground, grass). If $\Lambda(\bar{x}) < \eta$, then the entire sensor return did not come from background (i.e. obstacle, wall, fence, rock).

Structured Lighting

A structured light system was also developed and tested favorably under afternoon sunlight conditions. Our initial investigation used a laser line projected in front of a moving vehicle. Algorithms were developed to monitor the position of the line and to detect deflections of the line as the vehicle encounters obstacles. We also investigated the application of an array of simple spot lasers. Eye-safe spot lasers are more robust than a laser with line-generating optics in bright sunlight because all of the laser energy is concentrated into a small spot. The spot is much easier to detect than the line. Algorithms were developed to monitor the position of each spot as seen by a camera onboard the vehicle. The lasers and camera were arranged so that an obstacle would deflect the spot upwards as seen by the camera. The algorithm searched the epi-polar line corresponding to each spot, tracking the position of the spot from frame to frame. The presence of an obstacle is indicated by the movement of a bright spot up the epi-polar line at a rate that corresponds to current vehicle kinematics. The algorithm was extremely robust to false detections as the natural optical flow of terrain features was from top to bottom in the camera's field of view. The images below show how an obstacle deflects laser light such that an imaging system can detect its presence. Basic calculations identifying the obstacle's height, width, etc. were made from the physical geometry of the sensing system.

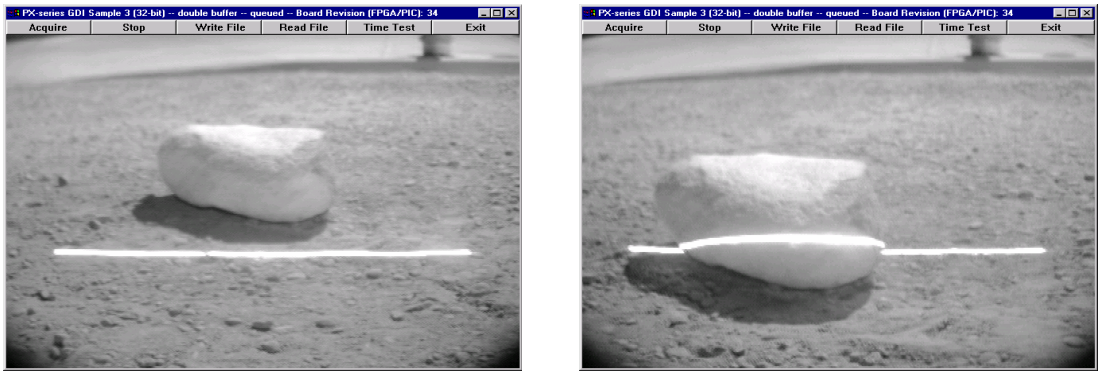


Figure 5. Obstacle deflecting laser light

Laser Range Scanners

We also investigated the data from two scanning lasers, an Acuity AccuRange 4000-LV line scanner and a SICK Laser Measurement System 200. A laser range scanner is an active scanning device that produces and projects laser light and makes range measurements based on the time of flight of the reflected laser energy. Data from both systems was taken in various outdoor environment scenes.

The Acuity AccuRange 4000-LV line scanner consists of a motor with encoder and a mirror mounted on the motor. The mirror deflects the outgoing beam from the sensor, and sweeps it through 360 degrees as the mirror rotates at a maximum rotational speed of 2600 R.P.M. The sensor and scanner are mounted on a flat plate that holds the mirror in the proper location relative to the sensor. The plate causes a blind spot of about 60 degrees of arc. Incoming data include temperature, pulse width range signal, ambient light signal and amplitude signal. The AccuRange High Speed Interface was used to increase the sample rate capability of the AccuRange sensor. The high speed interface board can sample the outputs at up to 50,000 times a second and buffer the data for reading by the host computer along with controlling the line scanner.

The Acuity line scanner was attached to the left front bumper of a truck. The data was stored on a PC/104 stack running a DOS program to interface to the high speed interface board. Figure 6 shows the field of view of the sensor. The rotating mirror was tilted so that the laser line swept the ground plane around the sensor, creating a

constant range value. This generated a conical field of view. The truck was driven on dirt roads, up to obstacles, and over rough terrain to collect data. The first scan was of a flat dirt road, shown in Figure 7. From the data, it can be seen that the sensor was not attached squarely to the truck, and therefore doesn't scribe a circular path centered around the sensor. The data from approximately 230 to 360 degrees is the truck. The noise on the left hand side is from the brush. A dirt road with ruts is shown in Figure 8. The right tire rut can be seen at 55 degrees, and the left tire rut can be seen at about 95 degrees. The embankment on the right hand side can also be seen as a decrease in the range on that side. Figure 9 shows the scan of a large bush off on the left as the truck is driven toward it. The scan of grass is shown in Figure 10. The underlying ground plane can be discerned from the chaotic range values caused by the varying heights of the grass and small scrub. A complex scene is shown in the final figure, Figure 11, a dirt road lined with scrub brush and trees. The differences in the range points are much greater when compared to the scans of the grass.

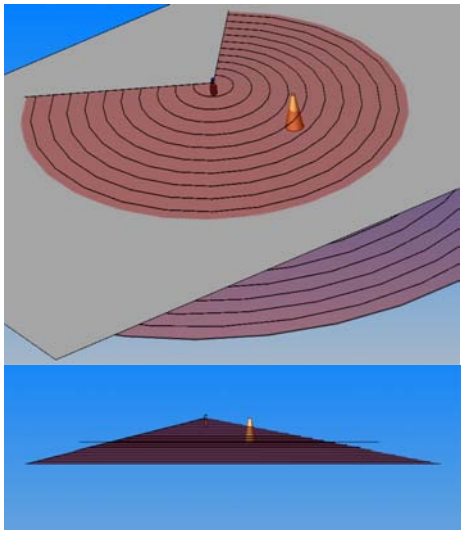


Figure 6. Acuity field of view

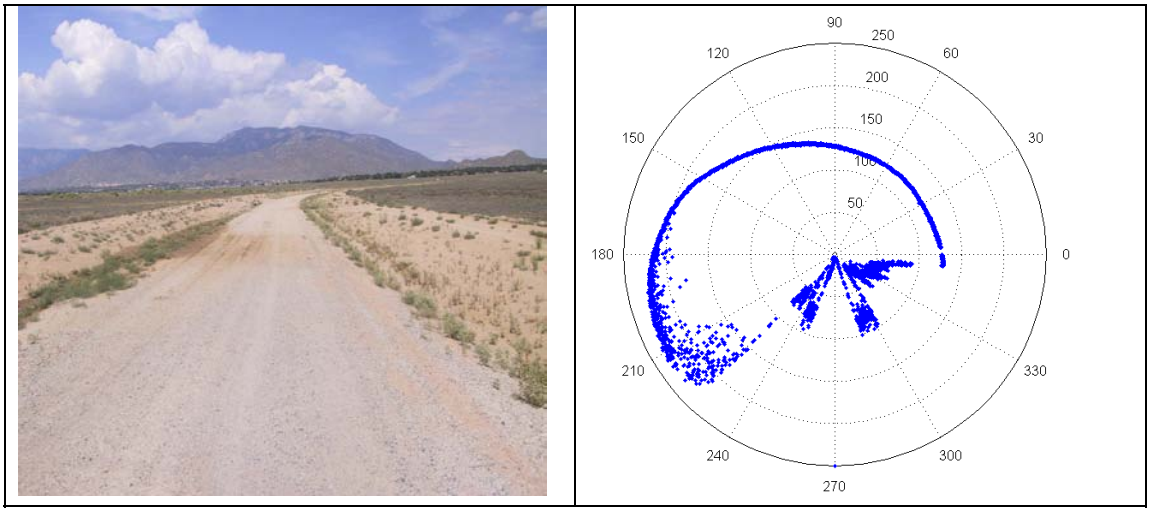


Figure 7. Dirt Road

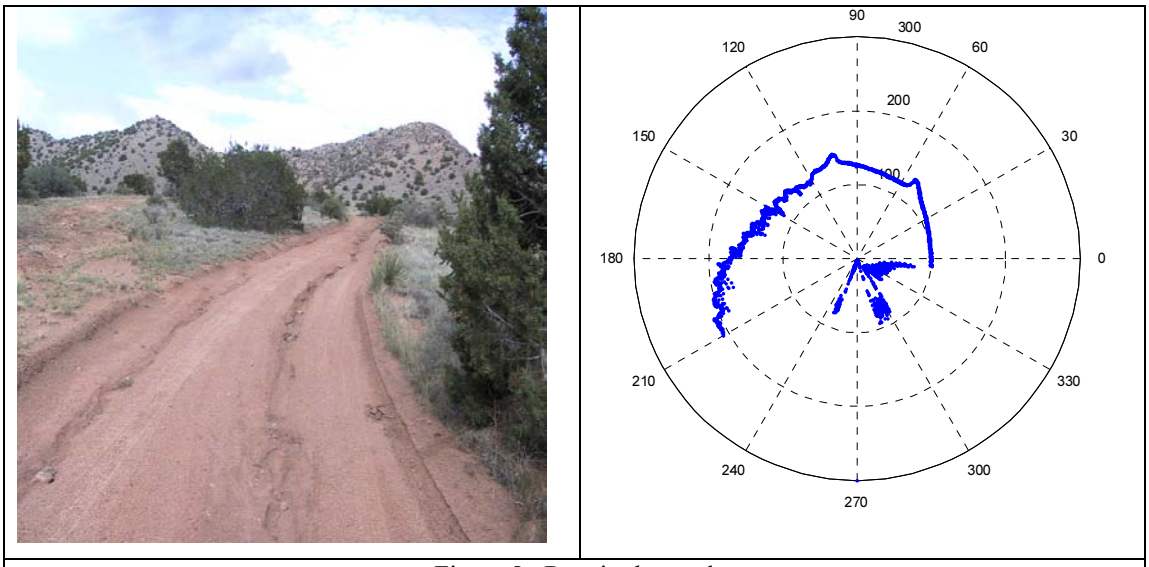


Figure 8. Ruts in the road

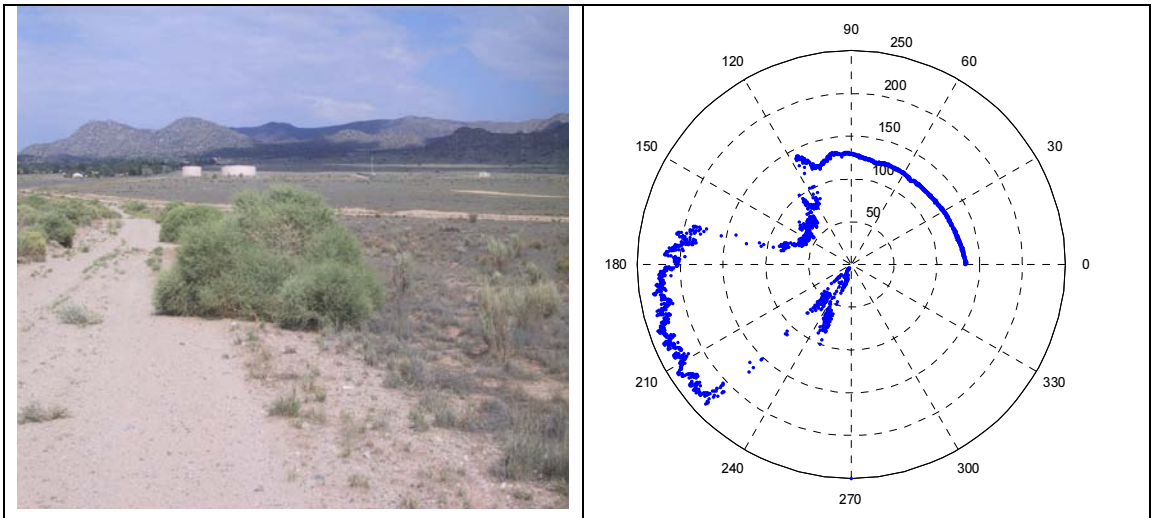


Figure 9. Bushes on the side of the road

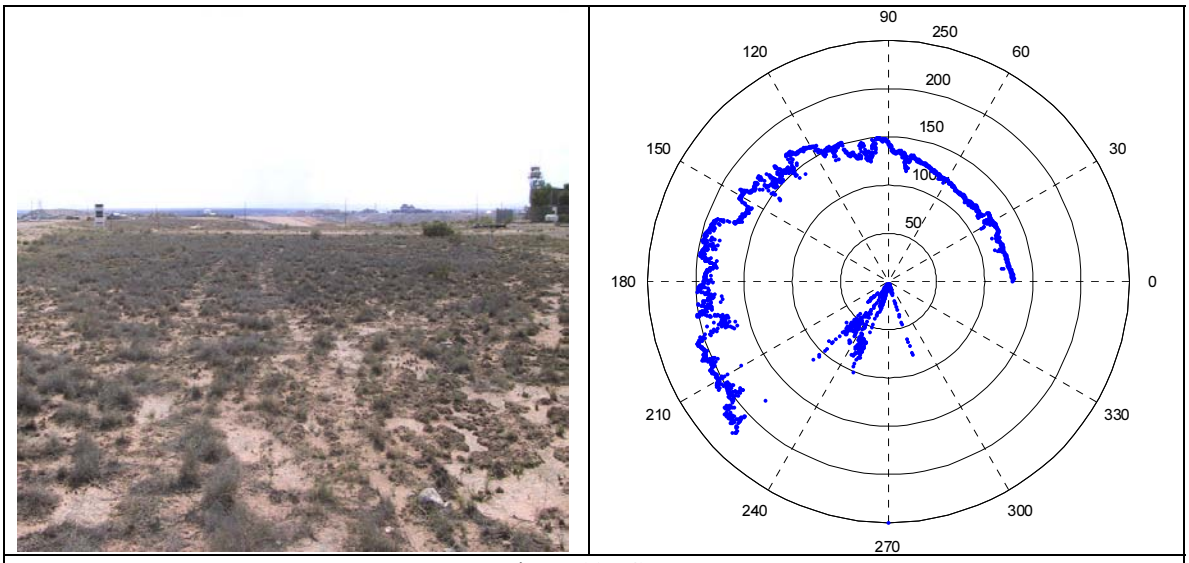


Figure 10. Grass

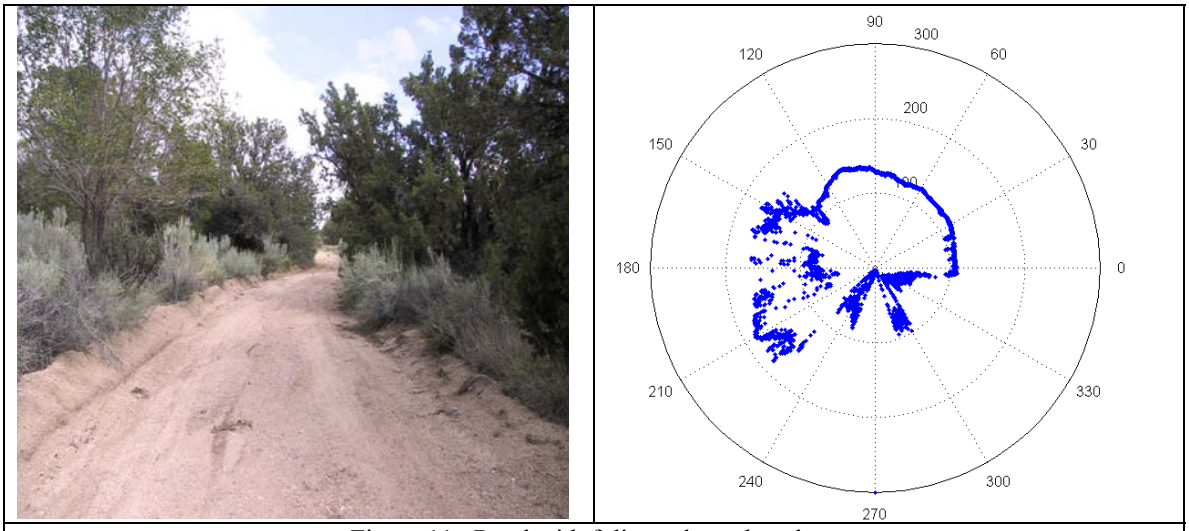


Figure 11. Road with foliage along the edges

The SICK Laser Measurement System (LMS) 200 emits a pulsed laser beam that is deflected through 180 degrees by an internal mirror. In the radial field of view, a light impulse (spot) is emitted every 0.25, 0.5 or 1 degree with a response time of 53 ms/26 ms/13 ms respectively.

The measurement data corresponding to the surrounding contour scanned by the device are given out in binary format via a RS 232/RS 422 interface.

The LMS was attached to a cart. The data was stored on a PC/104 stack running a Linux program to interface to the sensor. Figure 12 shows the field of view of the sensor. The LMS system was angled downward so that the fan sweep of the laser intersects the ground plane creating a planer fan swept field of view. The cart was pushed on dirt roads, up to obstacles and over rough terrain to collect the data.

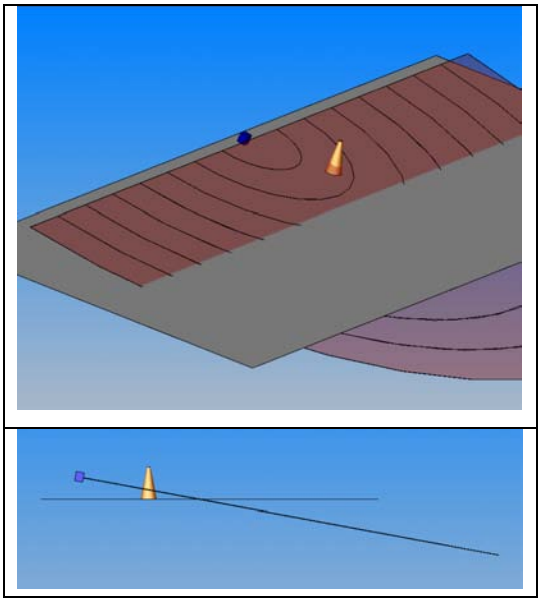


Figure 12. SICK LMS field of view

Similar to the data obtained from the Acuity AccuRange scanner, Figure 13 shows the scan of a flat dirt road. The road itself was not level, and therefore the data shows a slight tilt in the road between the angles of 40 to 130 degrees. The built up embankments off to the left and right can be seen. The dirt road is concave down and the data reproduces this clearly.

Figure 14 shows the scan of three cones placed in front of the LMS. Comparing the image to the data, the rightmost cone shows up at about 85 degrees, the middle cone at 91 degrees, and the leftmost at 149 degrees. The fourth obstacle on the scan is a small tree directly to the left of the sensor, outside the frame of the image.

The data as the sensor approaches a bush is shown in Figure 15. The line of points that run along the left side is from the fence that can be seen along the left side of the image. The bush, in the center, is shown as an obstacle with a fairly good form, but with some variation between the range points.

The sensor view of a fence is shown in Figure 16. As the sensor is moved toward the fence, the data is seen to divide into two distinct line segments. The first is from the ranges returned from the chain-link fence, the second is from the ground.

The final set, Figure 17, shows how grass is scanned. Unlike the data from the Acuity, the data from the SICK LMS bounces around in orientation since the cart bounced around as it was being pushed. Similar to the Acuity data, the underlying ground plane can be discerned from the chaotic range values caused by the varying heights of the grass and small scrub.

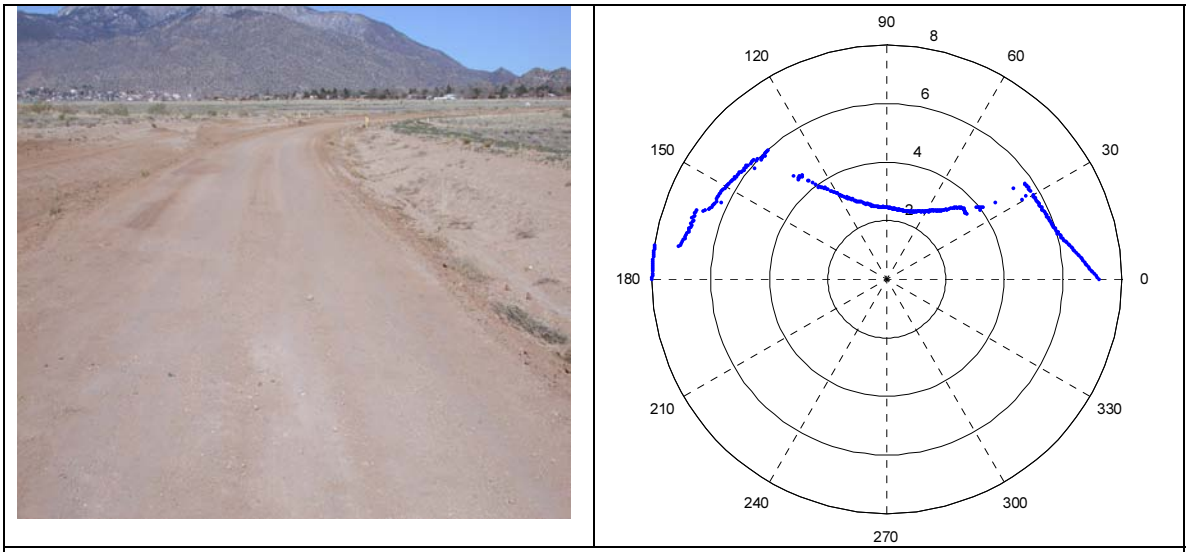


Figure 13. Dirt road

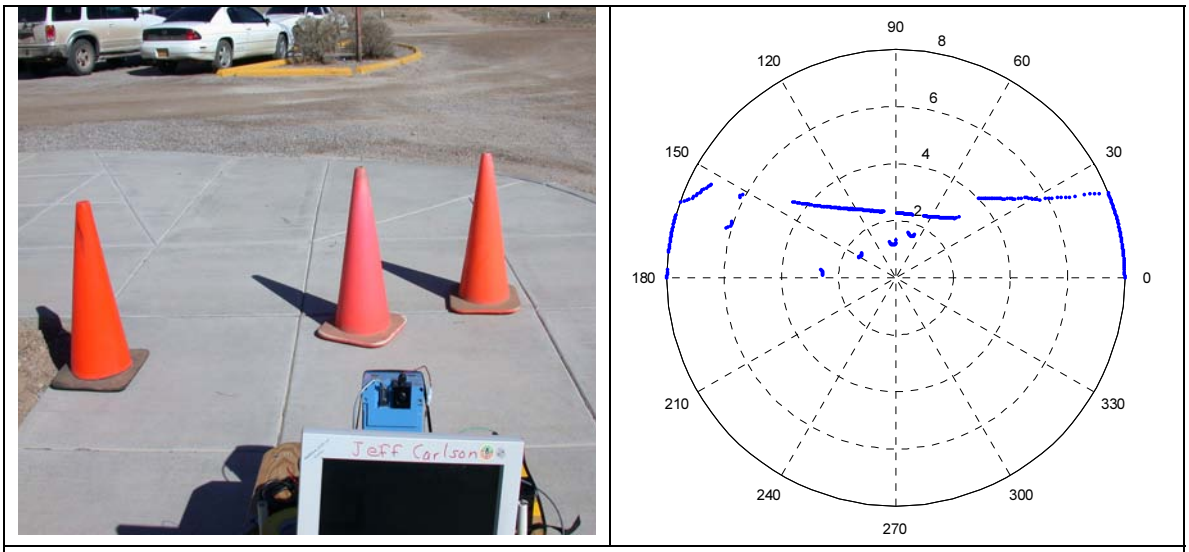


Figure 14. Cones

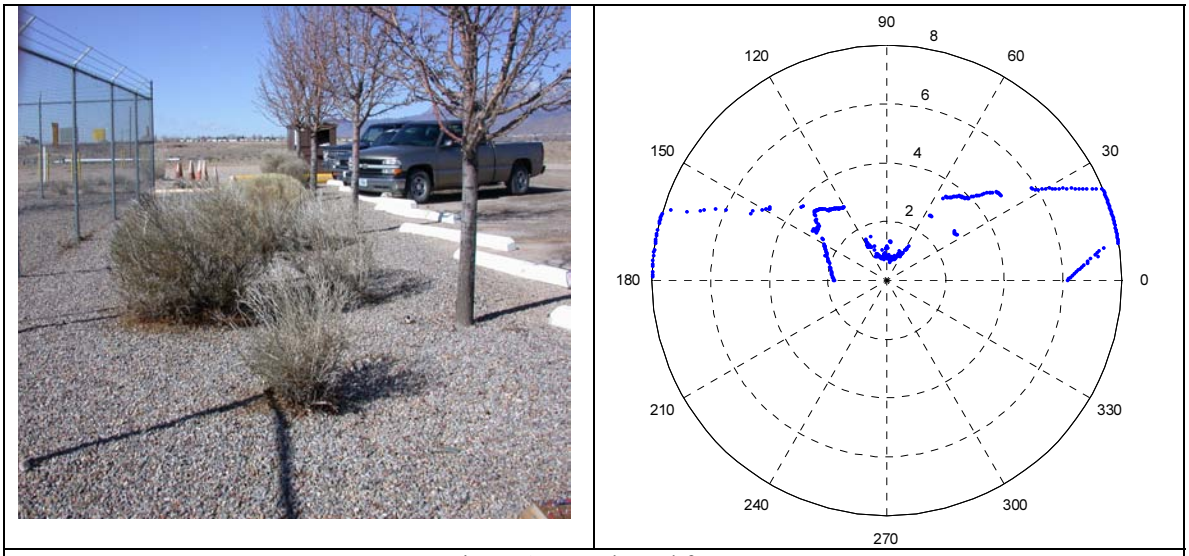


Figure 15. Bush and fence

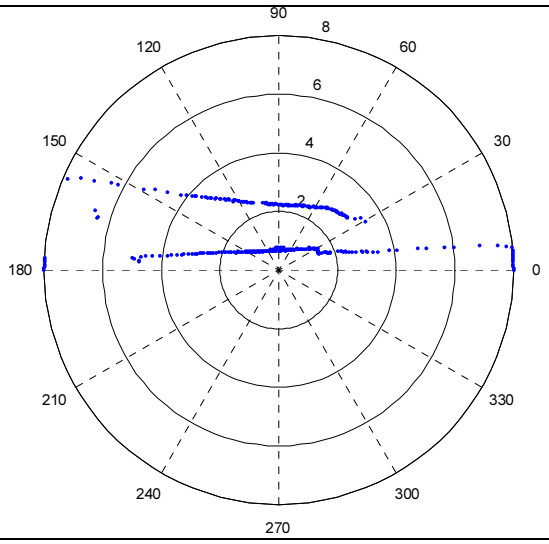


Figure 16: Fence

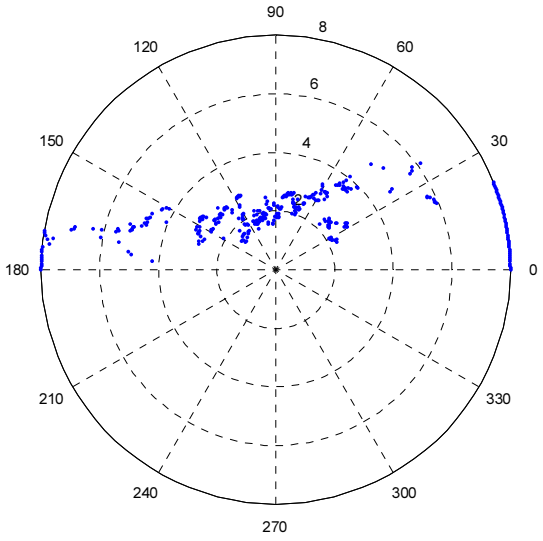


Figure17. Grass

Stereo Vision

A passive stereo vision sensor, called the Digiclops, was also investigated for obstacle detection purposes. The Digiclops software window shown in Figure 18 contains three children windows: an actual camera view (upper left), a disparity image (lower left), and a range image (center). The sensor provides real-time 3D information and has potential advantages compared to the active, laser-based systems. One advantage is the ability to measure the shape and volume of objects in the scene. One disadvantage is that night-time operations would require the use of artificial lighting, where as the laser-based systems naturally provide their own light source (i.e., the laser). However, the use of a passive system is beneficial in covert, day-time operations.

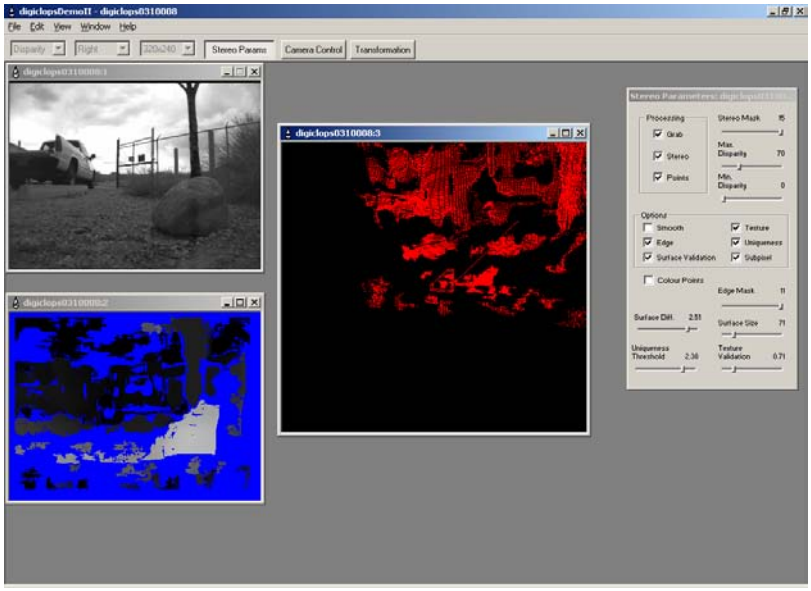


Figure 18. Digiclops software window

Scanner-less Range Imager

Sandia's Scanner-less Range Imager (SRI) was also investigated for obstacle detection purposes. Images of a typical bush-type obstacle and a rock are shown in the figures below. The SRI simultaneously provides both intensity and range images. Figure 19 is the intensity image as would be seen by a conventional black and white camera. Figure 20 is the range image. The SRI operates in real-time with sub-inch resolution. It's usefulness for obstacle detection from a mobile platform, however, is problematic. The range image is computed based on three consecutive frames. If the sensor platform is moving, these images are not registered and the range results are ambiguous. The images below were acquired from a stationary platform.

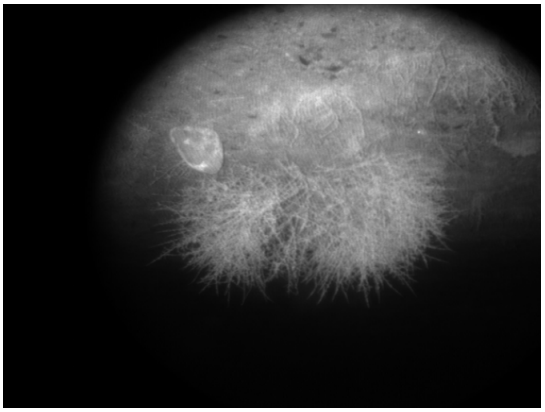


Figure 19. Intensity image

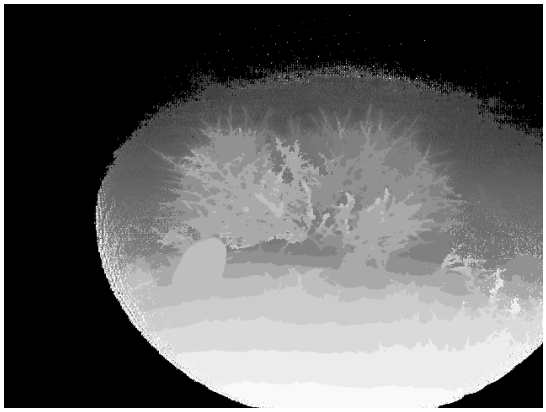


Figure 20. Range image

Laser Radar (LADAR)

We also evaluated the potential for obstacle detection in autonomous navigation applications using off-board sensors (i.e., not on the robot) using a state-of-the-art, high-resolution LADAR terrain mapping system. LADAR is a way to measure shapes in 3D based on laser illumination. Sensor data was provided by the Defense Advanced Research Projects Agency (DARPA). Statistical analyses showed the DARPA data sets were only sufficient for reliably detecting obstacles on the order of 3 meters in diameter. The median sample spacing was approximately 0.75

meters. Figure 21 is the LADAR terrain map from an approximately 150-meter by 150-meter wooded area. The green, blue, and purple spots are portions of the canopy. The brown spots are returns from the ground. The dark areas are shadows that are not seen by the LADAR. The LADAR was unable to detect the location of tree trunks – knowledge of which is crucial for autonomous navigation purposes.

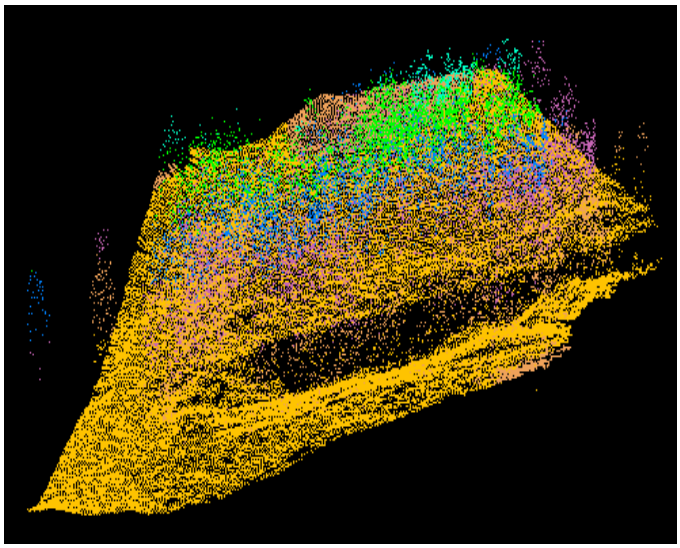


Figure 21. LADAR terrain map

Acoustic Imaging

Acoustic imaging provides many advantages for obstacle detection. These include its ability to sense obstacle depth (as opposed to estimating disparity with cameras), to infer obstacle density, to respond selectively to certain aggregate sizes (selective frequency), and to work in the absence of light. A vast majority of our effort was concentrated on researching acoustic imaging methods and image processing methods for obstacle detection. Three methods for acoustic image generation were considered; these included linear articulation of a sensor (scanning), acoustic lenses, and phased array imaging. In addition, image-processing methods are discussed for enhancement and interpretation of acoustic images.

Scanning

The simplest method of acoustic imaging is that of *scanning*. Scanning involves the physical articulation of an acoustic echolocation device with a narrow beam width. The device is articulated along the axis perpendicular to its line-of-sight. For each position on this axis, the echolocation device gathers range data such that the data can be applied in 2-dimensional space. If the echolocation device is articulated on two-axes, both perpendicular to its line-of-sight, the data can be applied to 3-dimensional space.

A simple 2-dimensional scanner was constructed in the lab using a Massasonic 5000/95 ultrasonic (95 kHz) sensor mounted to a linear stage. The linear stage was driven by a stepper-motor. The step counts were controlled using a function generator. Output from the Massasonic sensor was collected using an oscilloscope. Figure 22 shows a product photo of the Massasonic sensor and an oscilloscope screenshot of a typical output signal [1].

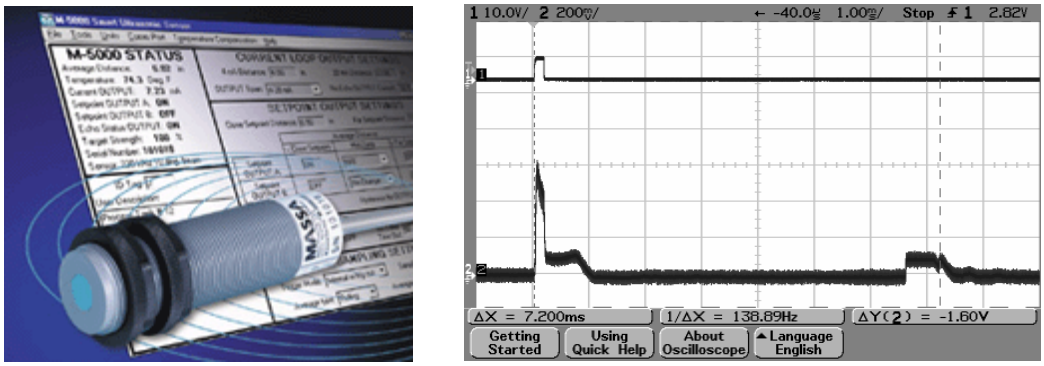


Figure 22. (a) Photo of echolocation sensor [1] (b) oscilloscope screenshot of output signal

Note from Figure 22(b) there are two waveforms. The top waveform shows a digital trigger indicating a transmission. In the second waveform, the initial peak indicates the point when the sensor is transmitting and the second peak indicates when the reflected sound is received. The data is easy to capture with an oscilloscope if triggered by the digital trigger. A linear stage with sensor attached was mounted onto a 4x6ft optical table, providing a large flat 2-dimensional space to place targets. An illustration of the system is given below in Figure 23 [2]. A photo of the setup is provided in Figure 24.

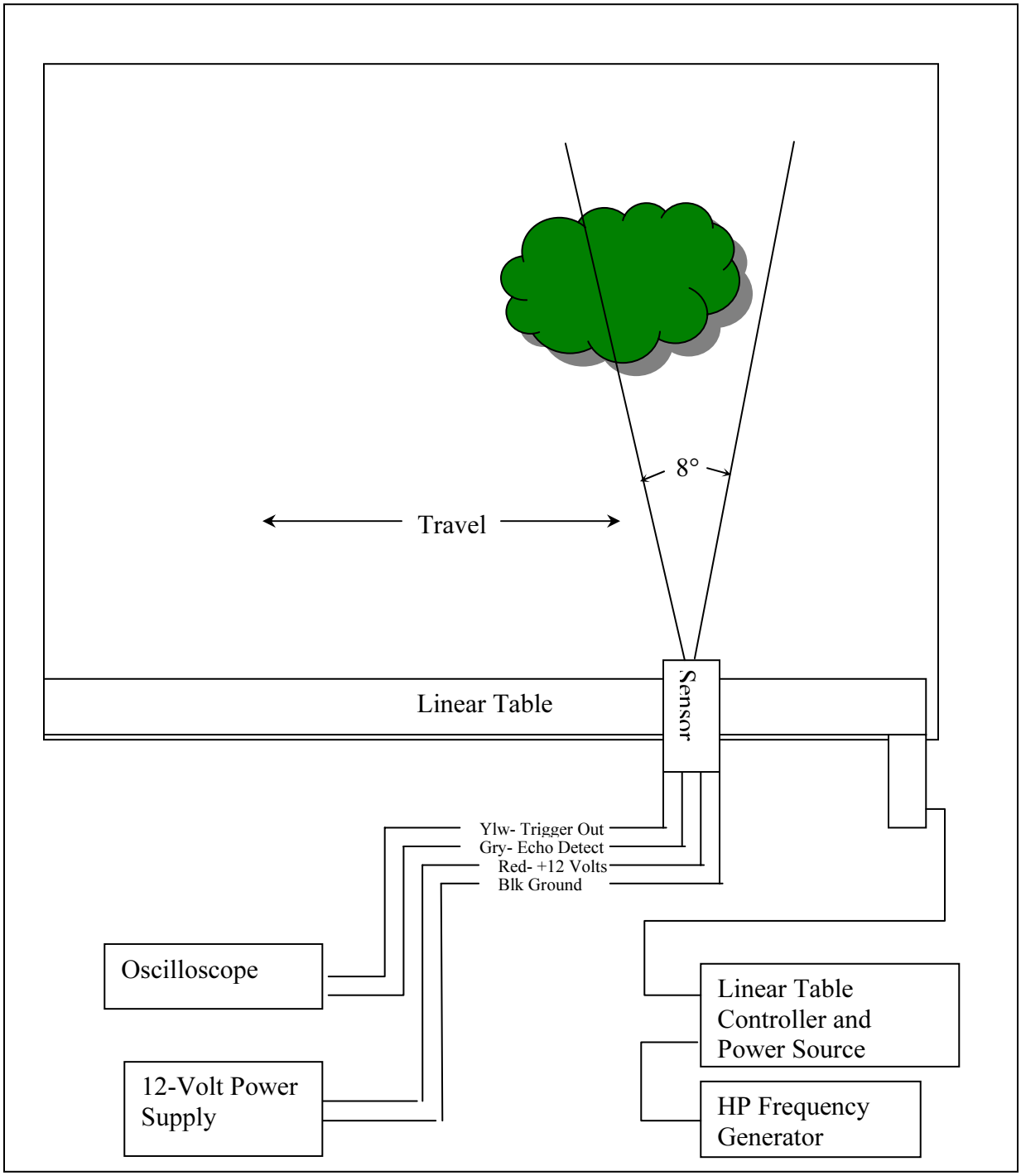


Figure 23. Illustrtion of scanner setup

Figure 24. Photo of echolocation sensor, stage, optics table and ancillary equipment

With the setup shown in Figure 24, several targets were positioned on the optics table and scanned, generating intensity data in 2-dimensional space. A simple MATLAB routine was developed to accept data files generated by the oscilloscope and combine these files side-by-side into an image [3]. Figure 25(b) below shows an example of an image produced (in MATLAB) through scanning of a rock and tumbleweed. Comparison with a photo of the setup suggests good results.

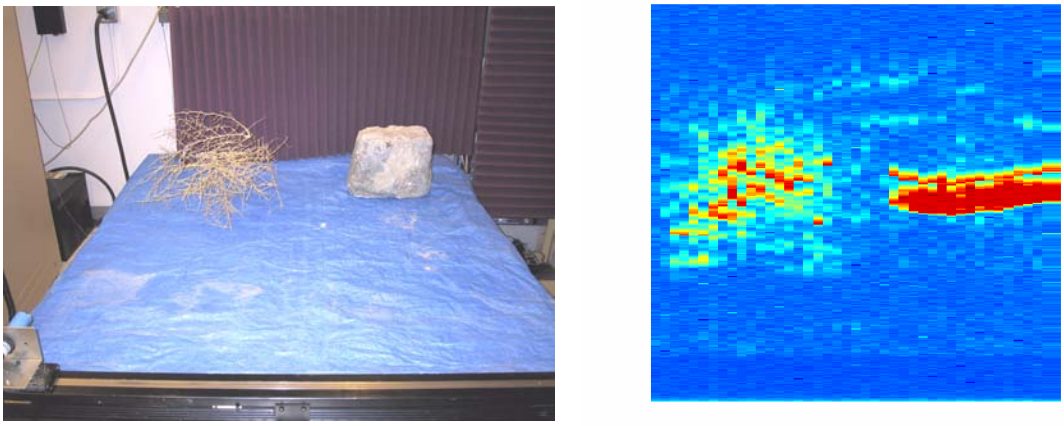


Figure 25. Shows (a) Photo of targets (b) scanner-produced image of targets

The tumbleweed shows up as a porous object, whereas the rock gives a solid return across its face. Note that the blue tarp in Figure 25(a) was added to protect the optics table from debris; the blue appearing in the background of Figure 25(b) is merely a coincidence. In addition, note that the image given in Figure 25(b) is pixilated more coarsely in the horizontal direction than in the vertical direction. The image is compiled of 43 separate datasets containing intensity values returned by the sensor. The spacing between scans was 1 inch; thus, the 43 scans span a 43-inch wide field-of-view. Each dataset contains 1000 intensity values sampled at 20 μsec intervals. The total time interval captured is thus 20msec, which translates into a 10-foot range. The scanned image shown was resized so that the same scale is applicable in both directions. Though the size of the tumbleweed is difficult to gauge, the size of the rock is simple to determine. It is seen that the image of the rock in Figure 25 (b) is larger than the actual rock. The image shows the rock to be 19 inches in width; the actual rock is only 12 inches in width. This discrepancy is attributed to the

angular resolution of the sensor. The published beam width value for the sensor is bw=8^{o} or bw=0.14^{R}. We thus state that the angular resolution is half the beam width as follows:

$$\Delta\theta = \frac{bw}{2} = 0.070^{R} \tag{1}$$

where $\Delta\theta$ is the angular resolution. From this quantity, we may define the spatial resolution. To begin, we assume that the sensor is articulated along the x-axis, and that all range data is gathered in the y-direction. Thus, the spatial (x) resolution becomes the following:

$$\Delta x = y\Delta\theta \tag{2}$$

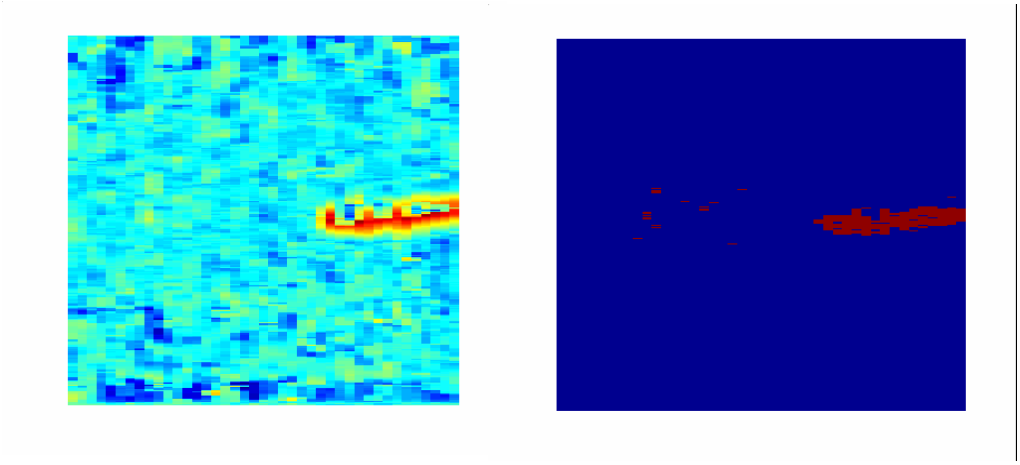
where Δx is the spatial resolution, and y is the distance of the target from the x-axis. For the setup shown in Figure 24, the face of the rock was positioned 4 feet or (48 inches) from the face of the echolocation sensor. Thus, the spatial resolution at the face of the rock is calculated to be the following:

$$\Delta x = y\Delta\theta = \left(48\,\frac{\text{inches}}{\text{radian}}\right)\left(0.070\ \text{radians}\right) = 3.36\ \text{inches} \tag{3}$$

and the system is said to have 3.36 inches spatial resolution at y=48 inches. If the spatial resolution value is counted twice, once for each edge of the rock, the width of the image of the rock would be 12 + 2(3.36) = 18.72 inches at y=48 inches. This is very close to the 19 inches of width seen for the rock in Figure 25(b). What is important to note here is that the spatial resolution represents an uncertainty value; this uncertainty grows larger as the target is moved further from the sensor. Another measure is the range resolution, Δy. The range resolution is dependent on the pulse width of the signal. One may be tempted to include other variables, such as subtle changes in the speed of sound, but these changes affect the accuracy of the range calculation, not the precision. This quantity can be difficult to control or anticipate. Given that only the rock face should be reflecting sound in the example above, the 'thickness' of the rock shown in the image is an indication of the range resolution. Simple observation indicates that the range resolution is several inches (~3-5 inches). Since the resolution is dependent on pulse width, a stronger reflection will result in a greater range resolution (greater range uncertainty).

Note also, that since the range uncertainty is caused by the pulse width, the uncertainty is always applied in the positive direction. More specifically, the exact location of a target is given by the temporal position of the leading edge of the output signal; any additional signal energy after that serves to suggest the existence of a target beyond (further way) than that point.

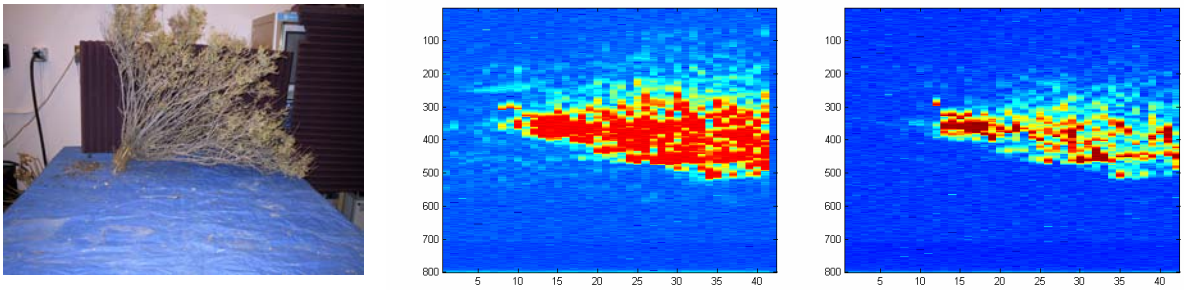
In order to distinguish between a high-burden obstacle such as the rock in Figure 25 and a low-burden obstacle, such as tumbleweed, one may apply several algorithms; two examples are presented here. Two methods are compared for the filtering of the image in Figure 25(b); Figure 26 shows the results of each. The first method uses a statistical filter; a window slides in the y direction, the values of the pixels within the window are histogramed, and the bin values in the histogram become features for a classifier. The classifier is trained for the tumbleweed (low-burden obstacle) and for the rock (high-burden obstacle); it is then used to generate an output image that represents only high-burden obstacles. This image is shown in Figure 26(a). The second method used simple thresholding, whereby a binary image is made through thresholding of the original image so that only 2% of the image area is displayed. Output from this algorithm is shown in Figure 26(b).



(a) (b)

Figure 26. High-burden obstacle detection using (a) Neural Network method (b) thresholding

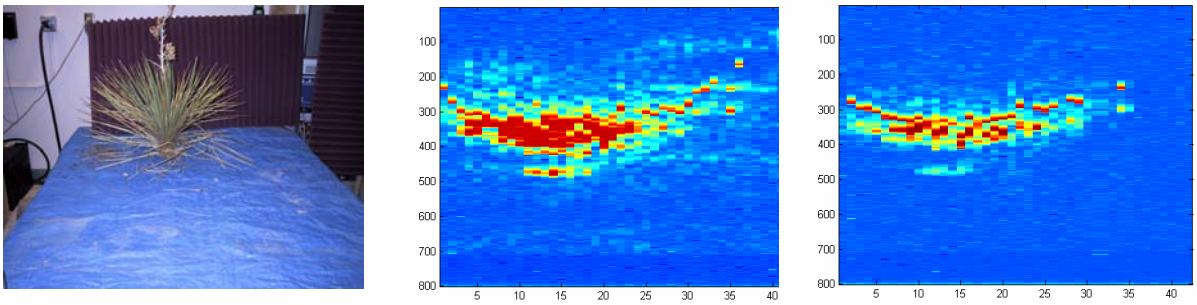
Other targets were chosen from an outdoor environment and tested to further investigate the scanning application for outdoor obstacles. In these next tests, we looked at some larger, more substantial bushes. Initially, the returns were very strong across the bush. It was not apparent that the obstacle was a bush or a rock. One problem is that the Massa sensor has a large, single gain; it does not allow the user to change this gain electronically.



(a) (b) (c)

Figure 27. Shows (a) photo of bush (b) image without foam (c) image with foam

The signal was thus attenuated using a 5/8th inch layer of open-cell foam placed directly over the face of the Massa sensor. This resulted in an image more "bush-like." Figure 27 shows the results of scanning a bush with and without this foam. Note that Figure 27(c) seems to discriminate the stem or stalk of the bush from the leafy portion. Another test suggesting some benefit for the use of foam to reduce gain was a scan of a Yucca plant. The results of this test appear in Figure 28 below.

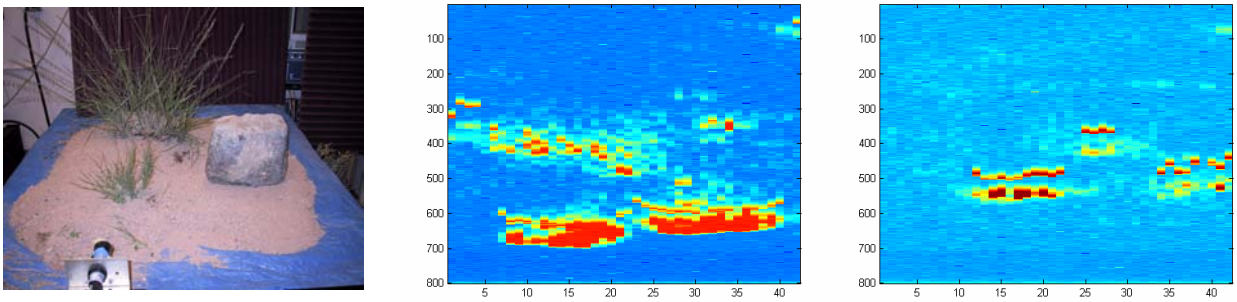


(a) (b) (c)

Figure 28. Shows (a) photo of yucca (b) image without foam (c) image with foam

For the yucca plant, one may see that the un-attenuated image (without foam) shows a solid density while the attenuated (with foam) image shows a more “bush-like” density.

Further tests include more than one object in the scene. The test has the same large rock as Figure 25, a small bunch of grass, and a larger bunch of tall weeds. The grass posed a problem. The wider flat blades of the grass, which were tightly interlaced, returned the same signal as the rock. Attenuating the signal with foam did not improve the image, such that the densities of objects were more discernable.



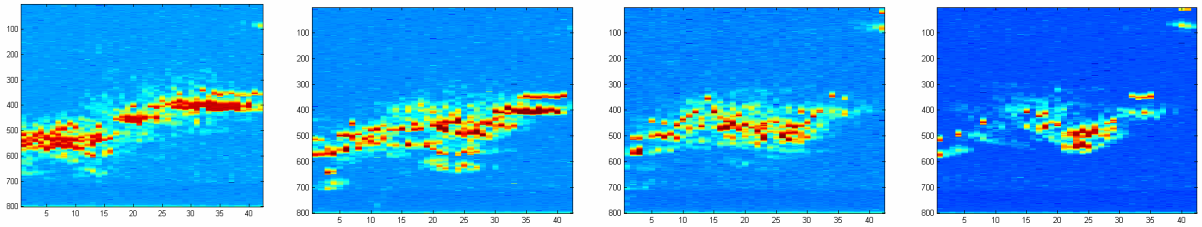
(a) (b) (c)

Figure 29. Shows (a) photo of scene (b) image without foam (c) image with foam

In order to add a vertical dimension to the test, a 4-stage flange was built, and the Massa sensor was used to scan a scene at four different heights. The flange appears in Figure 30(a), and the scene appears in Figure 30(b).

Figure 30. Photo of (a) 4-stage flange and (b) scene containing objects at different heights

The effort required the generation of 43 scans with the sensor mounted in each of the 4 mounting holes of the flange, thus creating 172 datasets. Representing the data was difficult, but it was decided that 4 separate images was the best method. These images are shown in Figure 31. By comparing the data taken at different heights with the photo in Figure 30(b), one may note that extra information is discernable in the image. For example, the rock pictured in the back right of Figure 30(b) is strongly discernable in Figure 31(a) and 31(b) but disappears in 31(c) and 31(d). The grass in the left foreground, however, is strongest in 31(a) but disperses away slowly with the progression of height. The bush in the center has a discernable stem in 31(a) then widens in 31(b) and 31(c), and begins to taper in 31(d).



(a) (b) (c) (d)

Figure 31. Images generated with sensor mounted at (a) bottom to (d) top of flange

The generation of acoustic images using a scanning method has many benefits and drawbacks. Its simplicity and predictability are apparent. In addition, the image generation itself is computationally uncomplicated. Unfortunately, however, since the sensor has such a narrow beam width, the generation of an image is likely to take some time. The experiments mentioned above required a man-in-the-loop procedure to take data, increment the stage, and/or reposition the sensor; so, data collection took several minutes. While an automated process would hasten the process, it would still likely take a minute or more to fully articulate the sensor take data with the precision shown above. Also important were the problems seen in discriminating objects. Figure 24 showed a clear distinction between obstacle and non-obstacle, but Figures 27 and 28 showed the need for gain control, and Figure 29 showed the ambiguities that still existed after gain control.

Acoustic Lenses

Acoustic Lenses were strongly considered in the acoustic imaging investigation. An acoustic lens focuses all acoustic energy from a particular angle onto one point. An "acoustic camera" may then be realized through the placement of transducers behind an acoustic lens such that each transducer isolates the acoustic energy received along a distinct angle. Figure 32 is provided to illustrate a crude example.

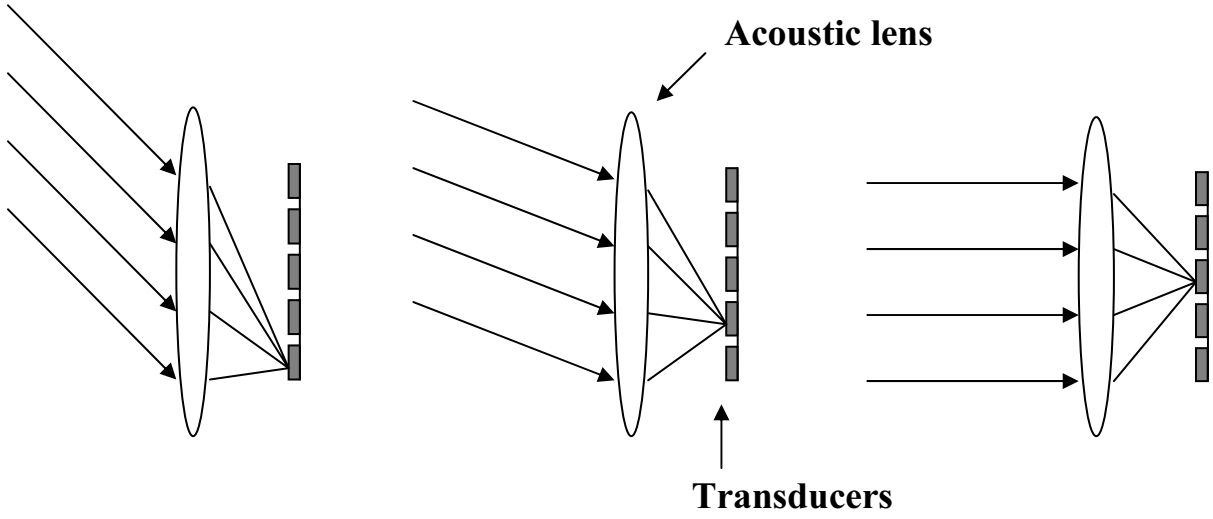


Figure 32. Shows the use of acoustic lenses to isolate acoustic energy along an angle

The use of acoustic lenses for imaging purposes is not a new idea. Acoustic lenses have been used much in underwater acoustic imaging applications. Figure 33 below shows the assembly of an acoustic camera, using two acoustic lenses, built and tested for underwater use at the University of Washington [4].

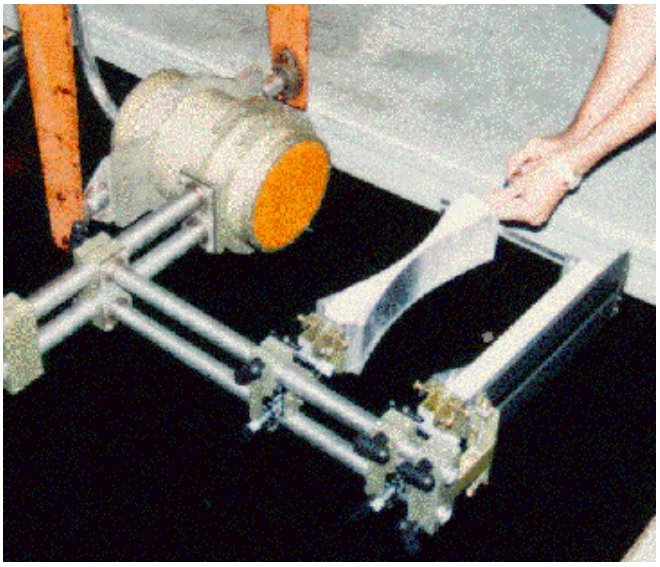


Figure 33. Photo of acoustic camera with two acoustic lenses

The use of acoustic lenses was investigated analytically and dismissed because of the physics-related limitations of coupling acoustic energy from air into any appropriate or constructible lens.

The challenge of acoustic lenses involves their use in air, particularly since the material properties of any lens will likely differ more between air and lens than between water and lens. Therefore, though acoustic lenses have been successful for use in water, the analysis to follow will demonstrate the improbability of constructing an acoustic lens that works in air. The analysis begins with simple geometric acoustics.

Geometric acoustics share some similarities with geometric optics; acoustic waves obey Snell's law, and have calculable transmission and reflection coefficients based on the properties of the lens and the surrounding medium. If we refer to the surrounding medium as medium 1 and to the lens as medium 2, then the model of the lens is reduced to two interactions: that between the 1-2 interface and then the 2-1 interface.

We begin by isolating the first interaction, the 1-2 interface. See Figure 34 below. Snell's law states that the angles of incidence and refraction are related by:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{c_1}{c_2} \tag{4}$$

where θ_1 and θ_2 are the angles of incidence and refraction in medium 1 and medium 2 respectively, and c_1 and c_2 are the speed of sound through each[3]. To calculate transmission and reflection, however, the medium density must be considered. The acoustic impedance of the material is given by the multiplication of the speed of sound through the medium and the density of the medium. This is given as follows:

$$Z_1 = c_1\rho_1 \tag{5}$$

$$Z_2 = c_2\rho_2 \tag{6}$$

where Z_1 and Z_2 are the acoustic impedances with units in *Rayles*, and ρ_1 and ρ_2 are the medium densities [4,5,6]. These acoustic impedances determine the amount of sound that is transmitted from one medium into the other and the amount of sound that is reflected. See Figure 34 below.

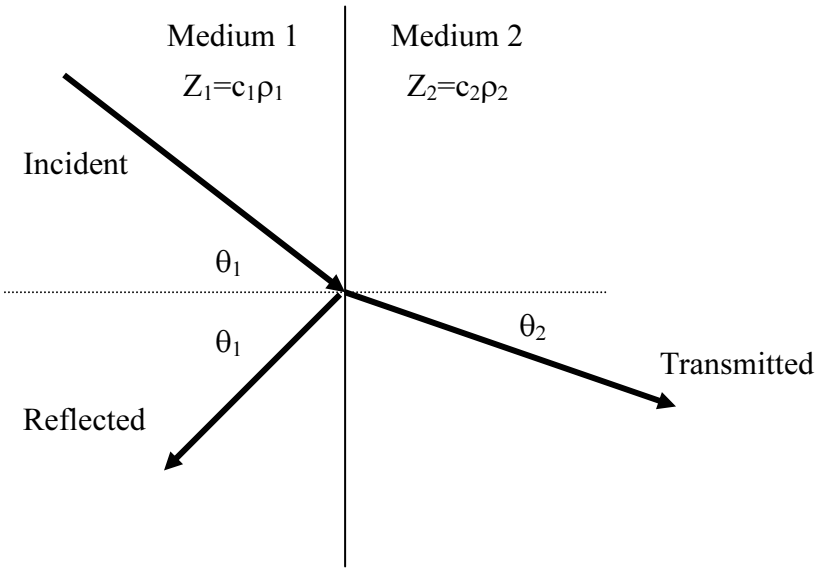


Figure 34. Illustrates refraction and reflection of an acoustic wave

To measure the amount of acoustic energy transmitted from medium 1 into medium 2, one must calculate the transmission coefficient. The transmission coefficient is the ratio of the acoustic wave energy of the transmitted wave in medium 2 to the energy of the incident wave in medium 1. If the angle of the incident wave is assumed to be zero, the transmission coefficient is expressed as follows:

$$T_{12} = \frac{4Z_1Z_2}{(Z_1 + Z_2)^2} \tag{7}$$

where T_{12} is the transmission coefficient for the 1-2 interface [7]. We may then calculate the reflection coefficient, the ratio of reflected to incident wave energies. We simply observe that any acoustic energy that is not transmitted is reflected. The result is the following expression:

$$R_{12} = 1 - T_{12} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \tag{8}$$

where R_{12} is the reflection coefficient for the 1-2 interface [7].

The expressions in (7) and (8) show the dependence of the material interface on using “like-materials.” If the lens material differs too greatly in density or in speed of sound from that of the surrounding medium, the transmission of sound will be greatly attenuated and the reflection of sound will dominate. To defeat this dependency, acoustic lens applications in water either use lenses that are in fact containers of liquid or constructed of solids with acoustic impedance close to that of water, such as polymethylpentene [8]. For liquid-filled lenses, a thin plastic or acrylic shell usually holds the shape of the acoustic lens. As long as the shell is much smaller than a wavelength and not too dense, the material properties of the liquid inside the shell will dictate the refractive properties of the lens [4]. Even this does not solve all problems, however, since temperature, pressure, and salinity can affect the speed of sound through surrounding water [4]. In fact, rather elaborate systems have been designed to maintain focal length in acoustic lenses despite environmental changes in the water around the lens. One system controls the temperature of the liquid inside a liquid-filled lens as a way to compensate for aberrations that result from changes in depth or surrounding water temperature, etc. [9].

Unfortunately, however, the use of acoustic lenses in air is not plausible given the expected transmission losses between air and any solid media. To illustrate the dependency, the relationships in (7) and (8) may be used to calculate the transmission losses for an air-lens interface. To be as realistic as possible, we attempt to find the most compatible solid material for use as a lens in air. Air has small acoustic impedance (Z_1 = 415 Rayles). The best material I could find information for was simple wood cork; it has the smallest acoustic impedance of all the solid materials I researched (Z_2 = 120,000 Rayles). The result of using a lens made of wood cork in air would be a transmission coefficient of T_{12} = 0.0137 (or –18.7 dB) and a reflection coefficient of R_{12} = 0.9863. At the far side of the lens, we get another attenuation of T_{21} = 0.0137 (another –18.7 dB) and reflection R_{21} = 0.9863. Given a total attenuation of –37.4 dB of the acoustic energy through the lens, not taking into account absorption, the use of the lens is dismissed.

The use of a solid material for construction of an acoustic lens in air is easily dismissed using the simple analysis above. Other ideas were considered, such as building an "air-lens" whereby the lens is made of a thin balloon-rubber that holds shape and contains an ideal gas [10]. Since the speed of sound through an ideal gas is dependent on gas density, there exist many possibilities (many gases) for controlled refraction. This construction, however, was not attempted. Another lensing method worth further research is the development of "sonic crystals." Emulating the effects of light refraction in crystalline structures, sonic crystals are structures constructed of cylinders (2D) or ball-bearings (3D) that have radii appropriate for the refraction of sound. Unfortunately, the structures described tend to be large (ie 1 meter long for 1.7kHz lens) [11]. Nonetheless, the idea may deserve consideration for future applications. A final consideration is not even lens-related but involves geometric acoustics just the same; a Schmitt-Camera, constructed of a parabolic mirror and an array of transducers at the focal length could be used to focus sound via reflection [10].

Phased Array (5-element Steered-Receive)

Perhaps the most well known application of phased array concepts is in radar. The operation of radar is dependent on wave-number transformations whereby the signals received from several precisely positioned antennae are phase-shifted and summed to isolate the direction from which the signal is received. The same principles may be applied to "acoustic radar" as well. By controlling the phase of sinusoidal acoustic signals simultaneously transmitted from transducers or speakers, a sound transmission may be steered or focused in space. Similarly, by delaying or phase shifting sinusoidal acoustic signals received by transducers or microphones, the reception may be isolated to sound directed from a particular direction or from a particular point in space (steered or focused receive).

Using wave-number relationships, a simple 5-element phased array imager, designed for steered-receive, was constructed. Testing of the device gave questionable results however. Using different phased-addition methods, a larger 16-element array, designed for focused receive, was constructed. Indoor and outdoor testing provided better results for this design. Both arrays were constructed using the Massa TR-89/B type 23, 23kHz 1inch diameter acoustic transducer [1].

For a steered-receive, phased-array approach, there exist four primary issues: gain, beam width, side-lobe angles, and the stand-off distance. Each of these is addressed in the design. The gain of the system is merely the ratio of electrical energy returned from the transducers and the amount of energy used to drive the speakers. The gain is dependent on several factors including: the efficiency of the speaker, the distance of the target, the reflection properties of the target, the efficiency of the receivers, the number of receivers, and the gain of the electronics used to drive the speaker and amplify the received signals. The gain of the system can be accounted for using an acoustic link-budget, a tool also suggested for radio communication. In a link budget, gains and signal strengths are converted to dB; they can then be added and subtracted based on the model. This is illustrated in Figure 35 below.

$$V_R = V_T + G_T - A_{FS} - A_M + G_R + G_{amp}$$

Figure 35. Receive signal calculated from source and known gains

One may find, for instance, that V_R is 10 dB short of the suggested receive signal strength at a given target distance, thus requiring the addition of some gain in the system (probably electronic). Unfortunately, values for the terms in the link budget are often difficult to determine. For example what is the gain for acoustic reflection off tree bark? Such terms must be measured comparatively, or gain must be added/subtracted purely by trial and error. For the 5-element phased array design, exact gains were ignored, and gain was simply observed to be proportional to the number of elements; a 5-element array has "5-element gain."

As was seen with the scanning system, beam width is an important parameter dictating the available resolution of the system. The side-lobes become important when one considers the field-of-view. For a linear phased array system performing a steered-receive, these two parameters are closely related. A linear phased array, with relevant terms, is illustrated in Figure 36 below.

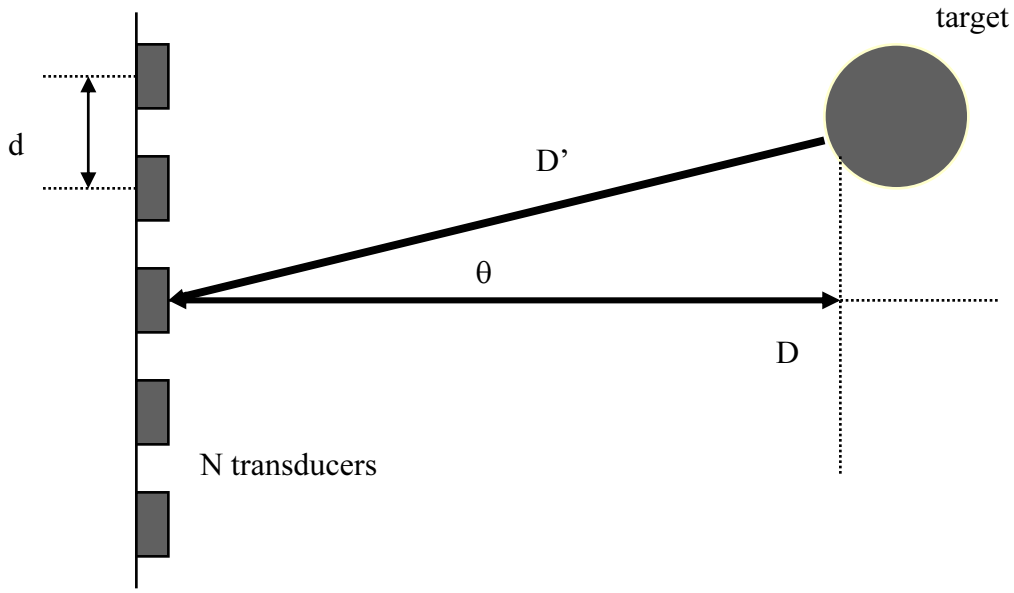


Figure 36. Illustrates the terms used for phased array analysis

For a phased array, the beam width is determined from the spacing of the elements *d*, the number of elements *N*, and the wavelength of the sound λ. The side-lobe angles are determined from the element spacing and wavelength. Both relationships are evident in the expression for intensity. By assuming that the target distance *D* is sufficiently large that θ becomes a small angle, and that the transducers are point receivers with zero diameters, the intensity of the summed signal is given by the following expression for multiple-slit interference:

$$I = I_0\left(\frac{\sin N\alpha}{\alpha}\right)^2$$
$$\alpha = \frac{\pi d \sin\theta}{\lambda} \tag{9}$$

where I is the intensity, I_0 is the intensity from one element, d is the distance between elements and θ is the angle to the array center (see Figure 36) [12,13]. By plotting I vs θ, one observes both the side-lobe angles and the beam width. A plot of I vs θ is given in Figure 37 for an array with d=1inch (minimum spacing for the TR-89/B), λ = ½ inch (23kHz), and N varied from 4 to 6 [14]. As noted above, the plot demonstrated that the side-lobe angles are not a function of N, though N affects the beam widths for each lobe.

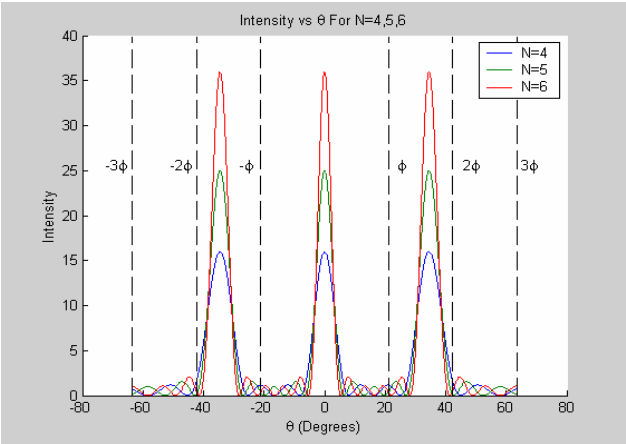
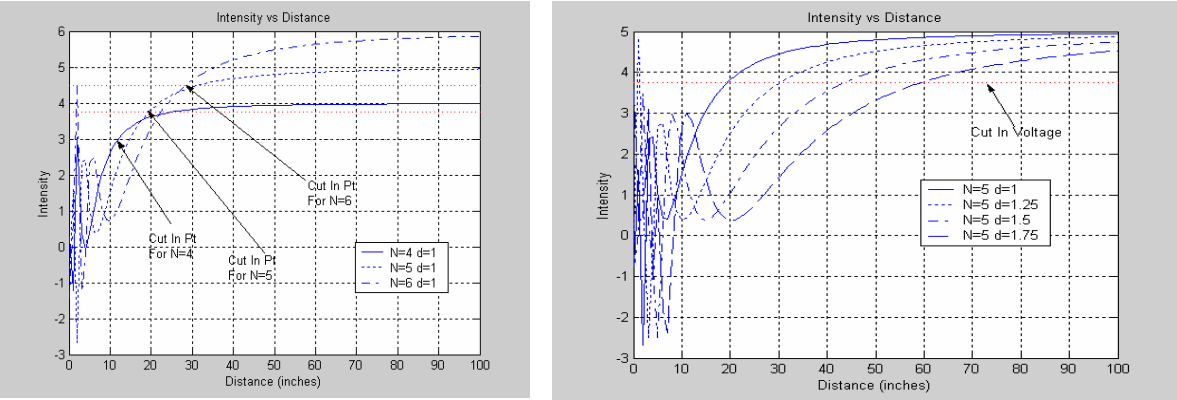


Figure 37. Plot of I vs θ given d=1inch and λ= ½ inch

Note that the center lobe at θ=0 is the maxima that will be steered using phased addition of the received signals. However, as the central lobe is steered, the side-lobes will rotate as well. The side-lobe positions dictate the allowable field-of-view. In Figure 37, the side-lobes occur at θ= -33° and θ= +33°. The field-of-view is thus limited to 33°, since any larger value would allow the array to interact with targets via the unintended side-lobes, generating ambiguities. The beam width is decided as the ½ intensity beam width and is easily determined from the plot.

A drawback of the steered-receive application is the requirement of a standoff distance. In order for equation (9) to be sufficient, the target must be far away enough for θ to be approximated as a small angle. To determine this distance, the analytical methods were used to start, but the results were difficult to interpret. Instead we determined the appropriate distance using a simple MATLAB simulation whereby an acoustic source with λ= ½ inch, is placed on the θ=0 line and moved further and further from the array (see Figure 38) [14]. Intensity I is evaluated for each distance D.



(a) (b)

Figure 38. Plots of normalized intensity I vs distance D given (a) fixed d (b) fixed N

It is noted from the plots in Figure 38 for close range (small D) the signals interact quite dynamically, with several maxima and minima being generated for small changes in D. The simulation ignores the spherical losses that would be expected with increased distance D; thus, after this dynamic region, the plot shows I asymptotically approaching N ("N-element gain") as the signals sum more and more in phase. It was decided arbitrarily that the sufficiently large distance would be the value for D whereby I=0.75N. This "cut-in" value, as is labeled in the plots, is seen to be a function of both N and d. This value for D was assumed sufficient and its effect on the I vs θ relationship was not investigated.

The design of the steered-receive imaging device requires only the determination of d and N to meet some criteria. The relationships in Figures 37 and 38 were both considered with respect to the desired results of the device. Since the device is intended for testing on top of an optics table with dimension 6x4 ft, the desirable "cut-in" distance would be at around D_{cutin} = 3 ft, so that the imager may view the far half of the table. The desired spatial resolution of the target was 4 inches for a distance of 6 feet. Given equation (3), this translates into an angular resolution (beamwidth) of $\Delta\theta = 3.18^{o}$. Other constraints exist as well; for example, the minimum value for d is 1inch since that is the radius of the transducer element. Unfortunately, values for d and N could not be determined to meet both the D_{cutin} and $\Delta\theta$ requirements. Figure 39 shows the allowable beam width both tabulated and plotted for each d and N [14].

	1 In	1.25 In	1.5 In	1.75 In	2 In
2	9.999	9.999	9.558	8.190	7.164
3	8.901	7.119	5.931	5.082	4.446
4	6.525	5.217	4.347	3.726	3.258
5	5.163	4.131	3.441	2.949	2.580
6	4.281	3.423	2.853	2.445	2.139
7	3.657	2.925	2.436	2.088	1.827
8	3.192	2.553	2.127	1.821	1.593
9	2.832	2.265	1.887	1.617	1.413
10	2.547	2.037	1.695	1.455	1.272

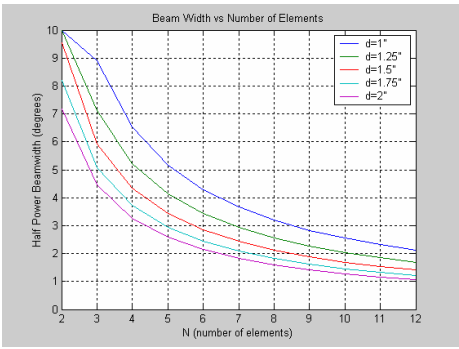


Figure 39. Shows beam width for different d and N combinations

It was decided instead that the design N=5, d=1 inch would be sufficient, essentially sacrificing $\Delta\theta$ (5.16°) for D_{cutin} (20 inches). This allowed the target to be sufficiently far at 36 inches where the spatial resolution would be Δx = 3.24 inches.

The device was built for N=5 and d=1 inch. The 5 transducers handle the reception of the reflected sound. The transmission of the sound was not steered and was thus achieved using a single high-power speaker. The completed assembly appears in Figure 40 [14].

To generate images, a signal generator was used to create 5 cycle voltage bursts at 23kHz to drive the speaker, sound bursts reflected off the target were received by the transducers, the signal from each transducer was captured using an oscilloscope and saved to data files. These files were then imported into MATLAB and summed with phase delays corresponding to different angles. Unfortunately, this process failed to produce a very good image for even the simplest tests. Figure 41 shows a setup of a large metal can being used as the target and its resulting image [14].

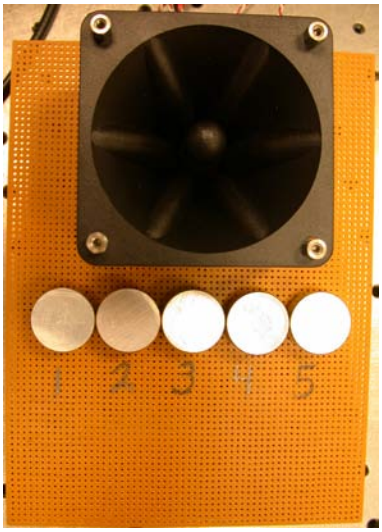
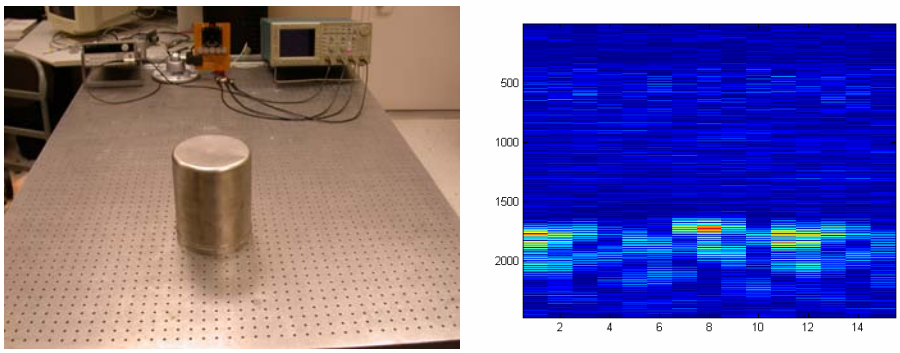


Figure 40. Photo of 5-element steered-receive phased array



(a) (b)

Figure 41. Shows (a) photo of setup with target (b) image of target

It was only determined much later that one possible problem with the above design was that the transducers were assumed to be point receivers with zero diameter. This assumption allowed the use of a multiple-slit interference model. However, since the transducer diameter was large, equal to the spacing in fact, it was determined that the model should include the effects of single-slit diffraction as well.

The expression for intensity combining the effects of multiple-slit interference and single-slit diffraction is given as follows:

$$I = I_o\left(\frac{\sin N\alpha}{\sin\alpha}\right)^2\left(\frac{\sin\beta}{\beta}\right)^2$$
$$\beta = \frac{\pi b \sin\theta}{\lambda} \tag{10}$$

where b is the diameter of the transducer [13]. When b is large, specifically $b \sim d$, diffraction greatly attenuates the intensity of the interference pattern. This attenuation may have resulted in a reduced contrast between minima and maxima for the image in Figure 41(b).

16-element Focused-Receive

The 16-element focused-receive acoustic imager differs in many respects from the 5-element steered-receive acoustic imager described above. Rather than summing time-domain sinusoidal signals, the 16-element imager sums time-domain intensity functions marking the reception of an acoustic burst. Each time-domain burst is rectified and filtered to generate the intensity function. An algorithm is then employed to build the image pixel by pixel, solving the transformation that maps the time domain intensities to x,y positions in 2-dimensional space. This method is thus implementing a focused-receive rather than a steered-receive. It may thus be said that this method is better suited for the near-field while the steered-receive is suited for far-field.

The algorithm to build the image is most easily stated as the following. For an '*N*' element linear array whereby the microphone data is rectified and filtered, the resulting image is constructed pixel by pixel whereby the intensity of a given pixel at *x,y* is the sum of microphone intensities at distances mapped to by x and y; this is shown below:

$$Pixel(x,y) = \sum_{n=1}^{N} Mic_n(d_n(x,y)) \tag{11}$$
$$d_n(x,y) = d_t + d'_n(x,y) \tag{12}$$

where *Pixel(x,y)* is the two dimensional gray-scale acoustic image, Mic_n is the n^{th} microphone intensity data vector, d_n is the roundtrip distance from transmitter to target at *x,y* to the n^{th} microphone, d_t is the distance between transmitter and target and d'_n is the distance between target and the n^{th} microphone. A focused-receive array with relevant terms is illustrated in Figure 42 on the next page.

The solution to equation 11 is most easily understood as a sum of intersecting ellipses [10]. Since the intensity data vector returned by Mic_n only gives intensity and distance, this data maps to an ellipse in *x,y* with the intensity given at that distance. *Pixel(x,y)* is the sum of these ellipses, producing a solution for the target in x,y where the ellipses sum to the greatest intensity. This function will be more discernable in the raw acoustic images presented later in this section.

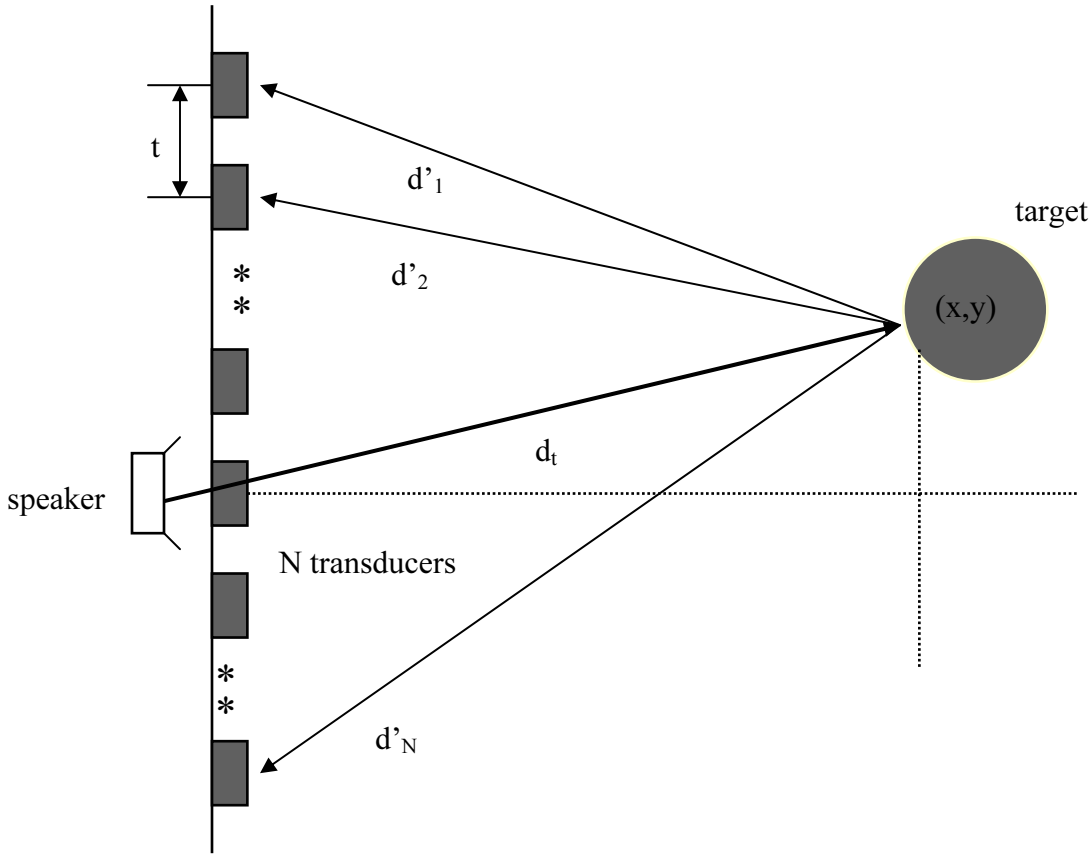


Figure 42. Illustrates focused-receive array with relevant terms

The motivation for this method involves the limitations of hardware. Specifically, it was desired that the application be performed on a PC104 computer. The PC104 has a 16-channel 40kHz A/D sampling capability. This did not meet the Nyquist minimum sampling criteria for the 23kHz transducers. In order to implement a phased array approach using the PC104 and for the purpose of investigating "something new," a different approach was taken to the phased addition of the received signals. Instead, the time-domain signal received from each transducer is amplified, rectified, and filtered using a simple circuit connected to each transducer. The resulting output is a simple time-domain intensity signal. When sampled, this signal becomes the intensity data vector described in equation 11.

One problem becomes very prevalent when using this method, however. When the sinusoidal signal is transformed into a mere intensity signal, the benefit of wave-number transformations may no longer be relied on to resolve the target position. This requires the intensity signal to be very narrow if the target is to be resolved with any precision.

To more easily address the problem of improving resolution, we focus on the reverse transformation, mapping from $d_1, d_2, \ldots, d_N$ to x,y. For the previous transformation of x,y to $d_1,d_2,\ldots,d_N$, the transformation is determined for each d value; given that the transmitter position is known, a given x and y value produce only one solution for each of $d_1,d_2,\ldots,d_N$ However, only two determined values from $d_1,d_2,\ldots,d_N$ are needed to solve for x and y. Using all of them seems "overkill." However, the motivation must be made clear.

Given that the image is generated using a recursive application of Equation 11, it is seen that the values for $d_1,d_2,\ldots,d_N$ are determined by their target x,y location. However, the intensities assigned to an x,y location are dependent upon intensity values found in every one of the N microphone data vectors. We thus focus on

the reverse transformation. By isolating the differentials for the reverse transformation, we may thus determine how errors in the distance (d) values reflect into errors in pixel locations. This is shown below:

$$\Delta x = \Delta d_1 \frac{\partial x}{\partial d_1} + \Delta d_2 \frac{\partial x}{\partial d_2} + \ldots + \Delta d_N \frac{\partial x}{\partial d_N} \tag{13}$$

$$\Delta y = \Delta d_1 \frac{\partial y}{\partial d_1} + \Delta d_2 \frac{\partial y}{\partial d_2} + \ldots + \Delta d_N \frac{\partial y}{\partial d_N} \tag{14}$$

where Δx and Δy are the differential placements or error in assignment of a pixel at x,y and Δd_n is the differential distance or distance error for the n^{th} element [15]. Δd_n is assumed to be roughly equivalent to the pulse width of a rectified and filtered acoustic pulse received by the n^{th} element after reflecting off of a target. Please note that Equations 13 and 14 attempt to simplify the resolution problem by discretizing the Δd_n values. The acoustic pulse is by no means square. Δd may be applied directly to a thresholded pulse, but in this case, it represents more of a *deviation*. Given that we wish to minimize Δx and Δy, for accurate location of the x,y pixel locations, it is apparent that small Δd values are ideal.

In addition to the dependence on Δd, equations 13 and 14 also illustrate the dependence on the partial derivatives. This dependence was thoroughly investigated in fact. Most of the analytical results were predictable however; for instance, it was showed that resolution improved (smaller Δx and Δy) when more transducers were used and the transducers were more closely spaced (large N small t). The most valuable contribution of this analysis, however, was to better illustrate the dependence of image resolution on target position. In other words, the partial derivative terms in equations 13 and 14 are themselves functions of x,y; this makes the resolution of a particular target dependent on the position of that target. The most important of these two resolution terms is the term representing the uncertainty of target width; this is Δx, the spatial resolution. The dependence of Δx on x,y can be best illustrated for a simple example. For a hypothetical system with N=3 transducers spaced t=20 units apart from one another, the resolution map for such a system is given by Figure 43. Figure 43 shows the 3 transducers placed at x=80, x=100, and x=120 (this is unitless). The resolution is indicated by color with *red* representing large Δx and *blue* representing low Δx. For this illustration, it is assumed that $\Delta d_1 = \Delta d_2 = \Delta d_3 = some\ constant$.

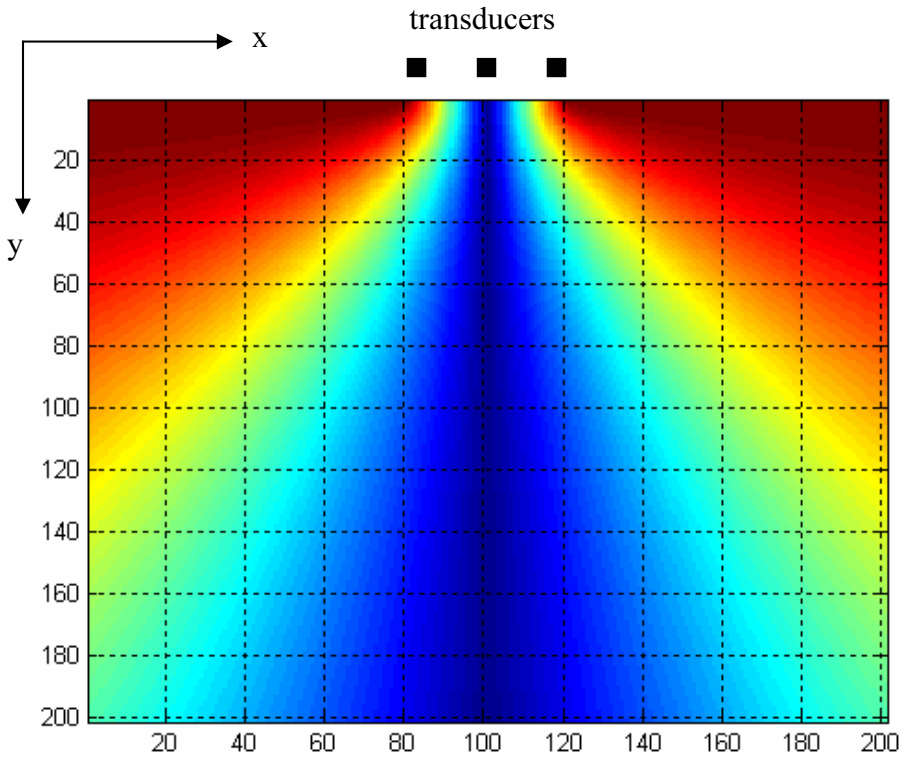


Figure 43. Resolution map showing Δx vs x and y

The illustration of Δx in Figure 43 is very encouraging. The resolution map suggests that the uncertainty Δx does not vary too greatly; in fact, it seems relatively constant (and small) over the centerline, suggesting that Δx does increase much with distance. This relationship is addressed with a cursory test presented later in this section (see Figure 57).

As stated for equation (11), the resolution issue may also be better understood as applied to a sum of intersecting ellipses. For an ellipse drawn in x,y for a given d_n, the widths of the ellipses are given by Δd_n.

A 16-element linear array was designed and built using the TR/89-B 23kHz transducers. Small receiver boards were created to mate with the transducer elements and condition the received signal for the PC104. Transmitter boards were created to drive high-frequency tweeters. Figures 44 and 45, on the next page, show the schematics for the transmitter and receiver boards respectively.

Each receiver board was designed to receive, amplify, rectify, and then filter the incoming acoustic pulse. The signal was then centered between 0 and 5 volts to be delivered to the A/D board on the PC/104.

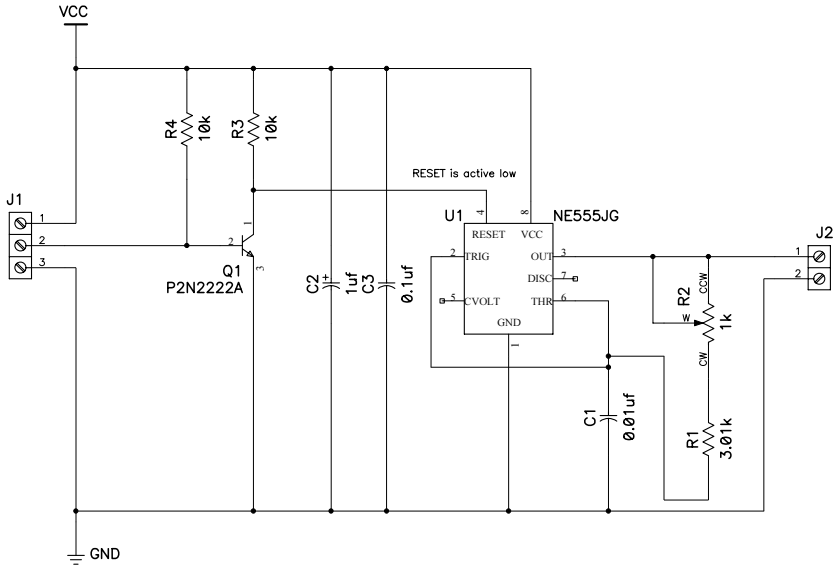


Figure 44. Schematic of Transmitter board

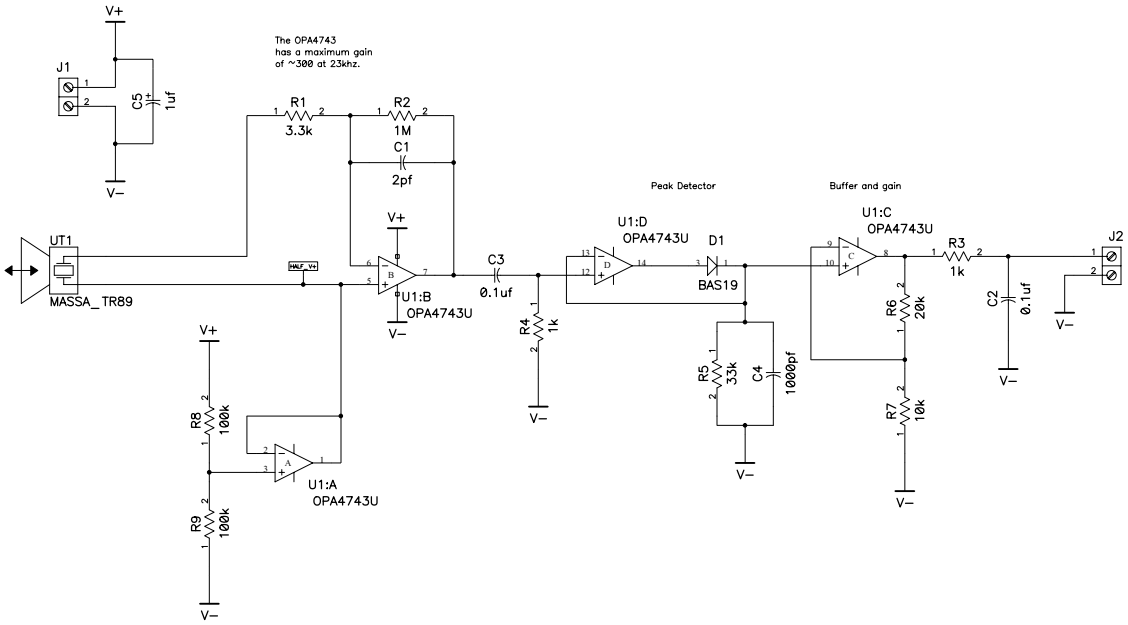
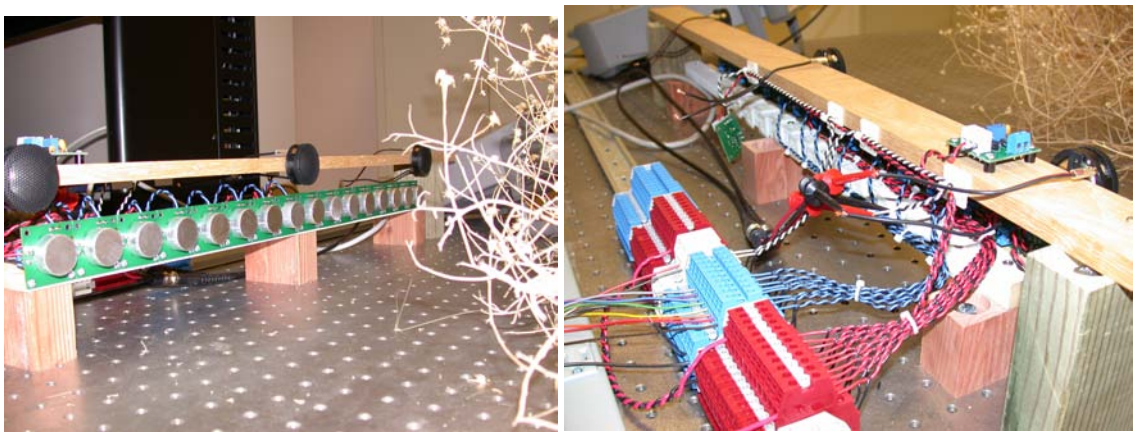


Figure 45. Schematic of receiver board

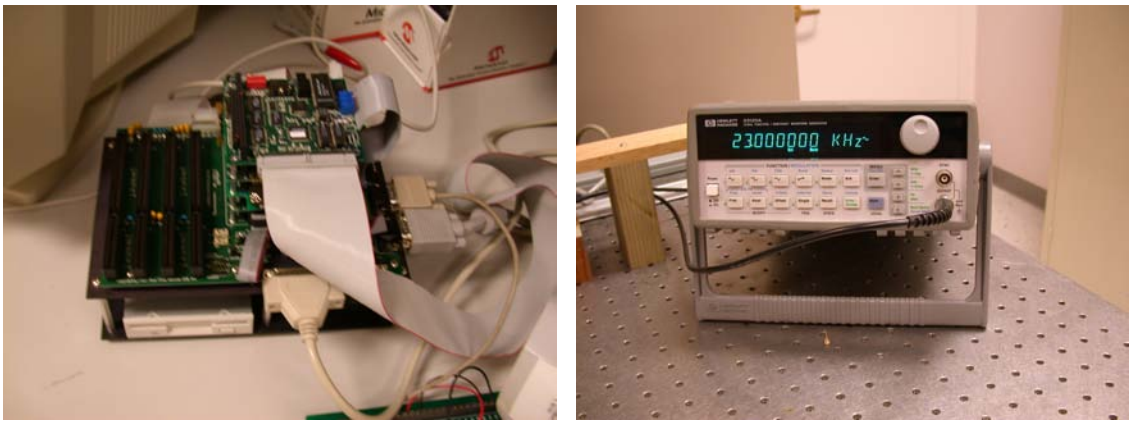
The completed receiver boards were 1½ inch square and were mounted edge-to-edge onto a wooden rail. This created a 24-inch long array with each of the elements spaced 1½ inch center to center. In actuality, the total length of the array would be better measured from the center of the first element to the center of the last element totaling 22½ inches long. Three tweeters were mounted to a wooden rail above the receiver rail. Though, three tweeters were desired, only two transmitter boards were fabricated because of inconsistent operation. These three tweeters were mounted in line with the edge of the first element, between the 8th and 9th elements, and then on the edge of the last element. Figure 46 shows the setup.



(a) (b)

Figure 46. 16-Mic 3-Spkr array shown from (a) the front (b) the back

Since the transmitter boards did not always work as desired, a function generator was programmed to generate a 1-cycle 23kHz acoustic pulse when its external gate was triggered. A PC 104 stack was assembled and programmed via an executable C++ program to trigger each of the three tweeters, collect the acoustic data from the receivers on an A/D board, and write this data to a file. Figure 47 shows the PC 104 stack and Function generator.



(a) (b)

Figure 47. (a) PC/104 used for data acquisition (b) Function generator used to drive tweeters

To generate an acoustic image, the data acquisition program needed to be run on the PC 104. This program would collect the data and then write it to a data file 'image.dat.' This data file was then loaded into a Matlab shell, and custom Matlab routines were used to generate an image through application of equation 11.

The receiver, with transducer and conditioning circuit, was tested by reflecting an acoustic pulse off a target and acquiring the signal with an oscilloscope. Unfortunately, this test showed that the pulse width of the signal produced by the receiver circuit was very large as is seen by Figure 48.

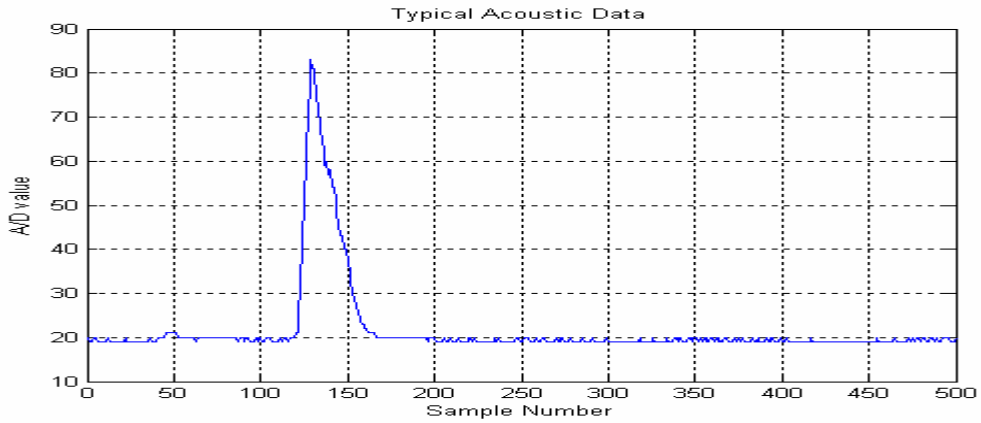


Figure 48. Measured acoustic pulse

Given that the A/D was sampling at 40kHz for this example, the pulse width is approximately 1 millisecond at the base or about a foot given the speed of sound in air.

Given the nature of the received pulse, it is easy to reason that the leading edge of the pulse represents the true distance. Unfortunately, the pulse distributes its energy over a foot with the peak energy shown near the center, about 3 inches from the true position. Given equations 11-13, the image would not be precisely represented with such large values for Δd.

It was thus decided that the acoustic pulse should be processed via a "*leading-edge detector.*" This was done in MATLAB since it was not built into the circuit.

The leading edge detector was implemented by a discrete filter set up to perform a 4-point derivative; the resulting positive values represent leading edges. Any negative values represented falling edges. Equation 15 below shows how this is done.

$$Mic\_leading\_edge(k) = -2Mic(k) - Mic(k+1) + Mic(k+2) + 2Mic(k+3) \quad (15)$$

Figure 49 shows Equation 15 applied to the signal in Figure 48.

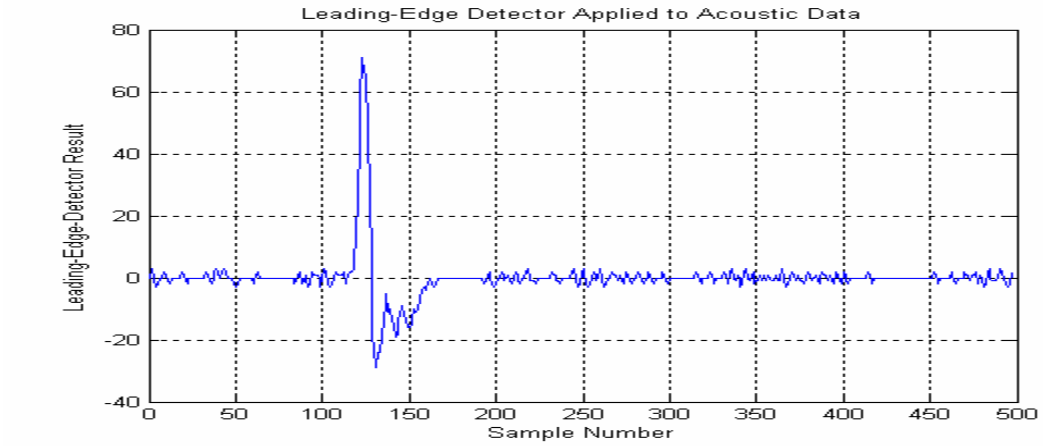


Figure 49. Leading-edge detector applied to acoustic pulse

By inspection of Figure 49, it is seen that the "*leading-edge detector*" does in fact focus the energy more correctly with greater accuracy and less "smearing." The negative edge is ignored.

To demonstrate the capabilities of the acoustic imaging system, obstacle scenarios were setup and imaged. For the first test, a thick metal bar is hidden inside a tumbleweed and a rock is placed behind the tumbleweed. The data acquisition program was run and a data file was generated. This data file was loaded into a Matlab shell. Matlab routines were run to initialize and describe the array setup, to apply the leading-edge detection to the data, to recursively apply Equation 11 for the generation of the base image, and finally to threshold the image. Figure 50 shows the complete setup from the side. Figure 51 shows a photo from the top along with the raw acoustic image and the thresholded image; note that the scales along the bottom and side are in *centimeters.*

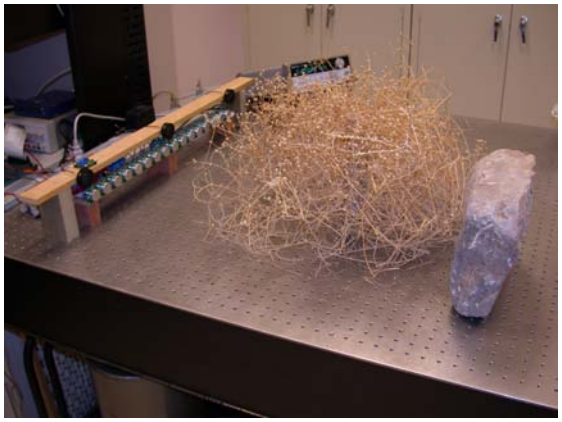
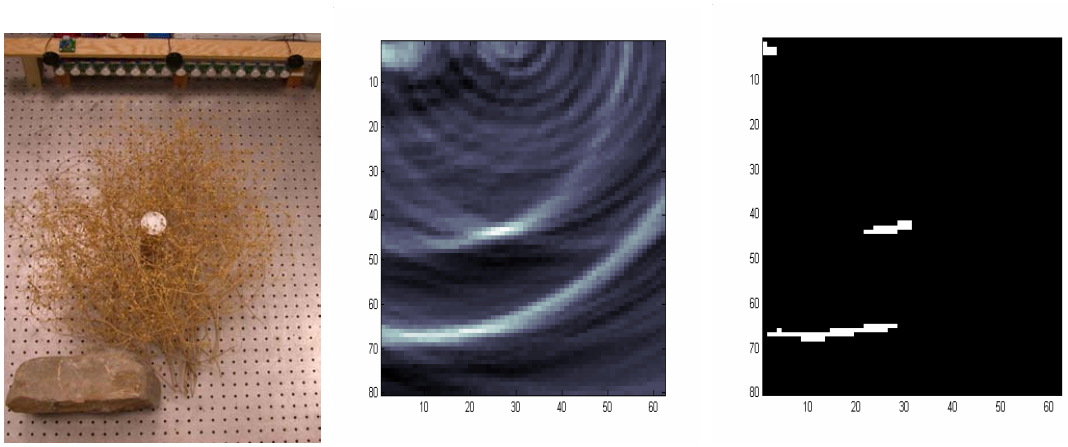


Figure 50. Setup with tumbleweed hiding a steel bar and rock

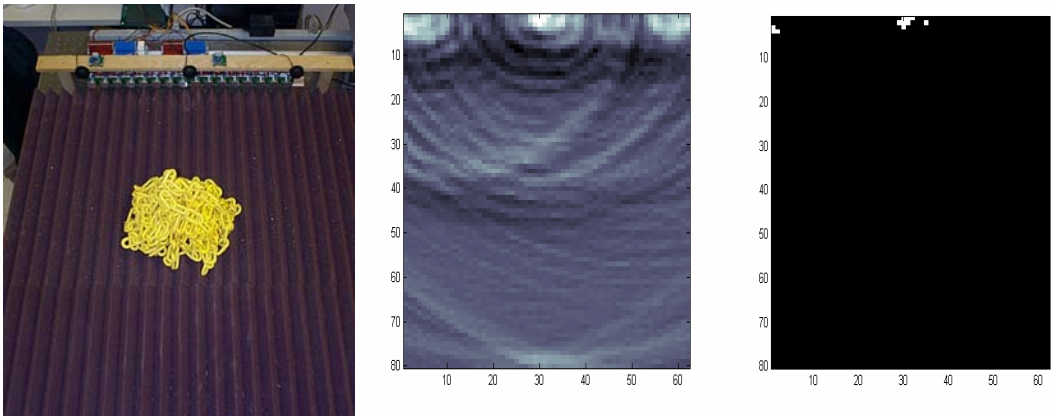


(a) (b) (c)

Figure 51. Shows (a) Top view photo (b) raw acoustic image (c) thresholded binary image

By examination of the acoustic image in Figure 51(b), one observes that the thresholded image shows clearly the location and approximate size of each dense obstacle while ignoring the presence of the tumbleweed. This result occurs because the operating frequency is 23kHz; so, the signal wavelength (about ½“) is much larger than the stems on the tumbleweed. This result is very different than what was seen with the scanner. For the scanning acoustic imager (95kHz), image processing was required to isolate the rock from the image. With this system, the lower frequency (23kHz) automatically isolates the high-burden obstacles in Figure 51. This effect, however, may be misleading, as is seen in the next example. For this test, a plastic chain is piled in front of the acoustic imager. Figures 52 and 53 show the setup and experimental results. One may reason that the height of the pile of chain would suggest it to be an obstacle. The purple foam seen in Figure 52 is acoustic insulation; it was put down after discovering that the holes at the far edge of the optics table reflected back sound.

Figure 52. Photo showing pile of plastic chain



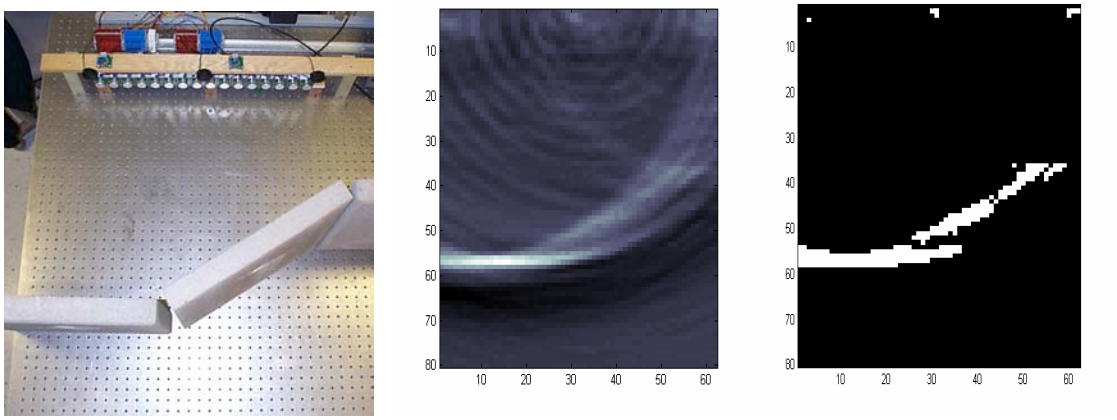
(a) (b) (c)

Figure 53. Shows (a) Top view photo (b) raw acoustic image (c) thresholded binary image

Unfortunately, the piled plastic chain does not appear at all in the binary image. It is believed that this is a result of the 23kHz wavelength (about ½") being larger than the diameter of a chain link (less than ¼"). It is not believed that a metal chain with similar link diameter would return a better result, though this assumption has not been verified.

The next test was setup to investigate the imager's ability to image walls or other flat obstacles. Concrete patio blocks were placed upright to produce the flat obstacles. They were placed both parallel to the imager axis and at a slant. Figures 54 and 55 show the setup and results of the experiment.

Figure 54. Photo of patio block setup



(a) (b) (c)

Figure 55. Shows (a) Top view photo (b) raw acoustic image (c) thresholded binary image

By observation of Figure 55(b), it is clear that the slanted patio block does not reflect as much acoustic energy back to the array. However, the binary image in Figure 55(c) still represents the correct position and orientation of both patio blocks.

To demonstrate the imager response to target position (see Figure 43), a simple target (concrete cylinder) was imaged in several different positions. The concrete cylinder is pictured in Figure 56 below. The test results for each target position are shown together in Figure 57.

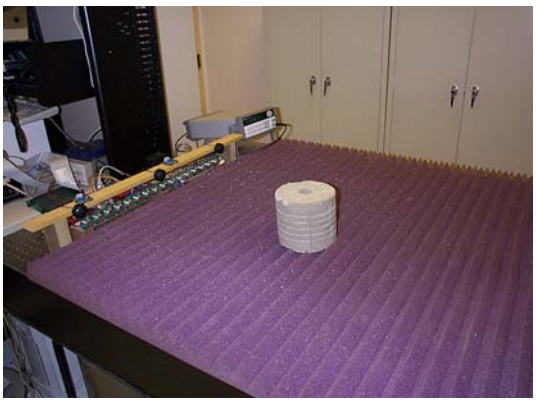
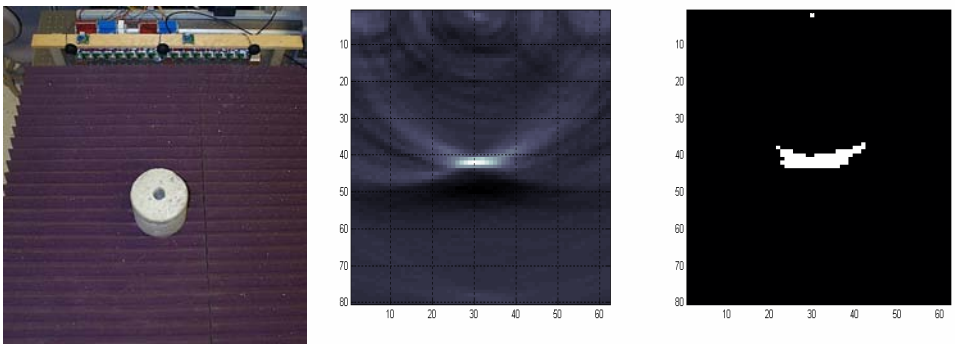


Figure 56. Photo of concrete cylinder



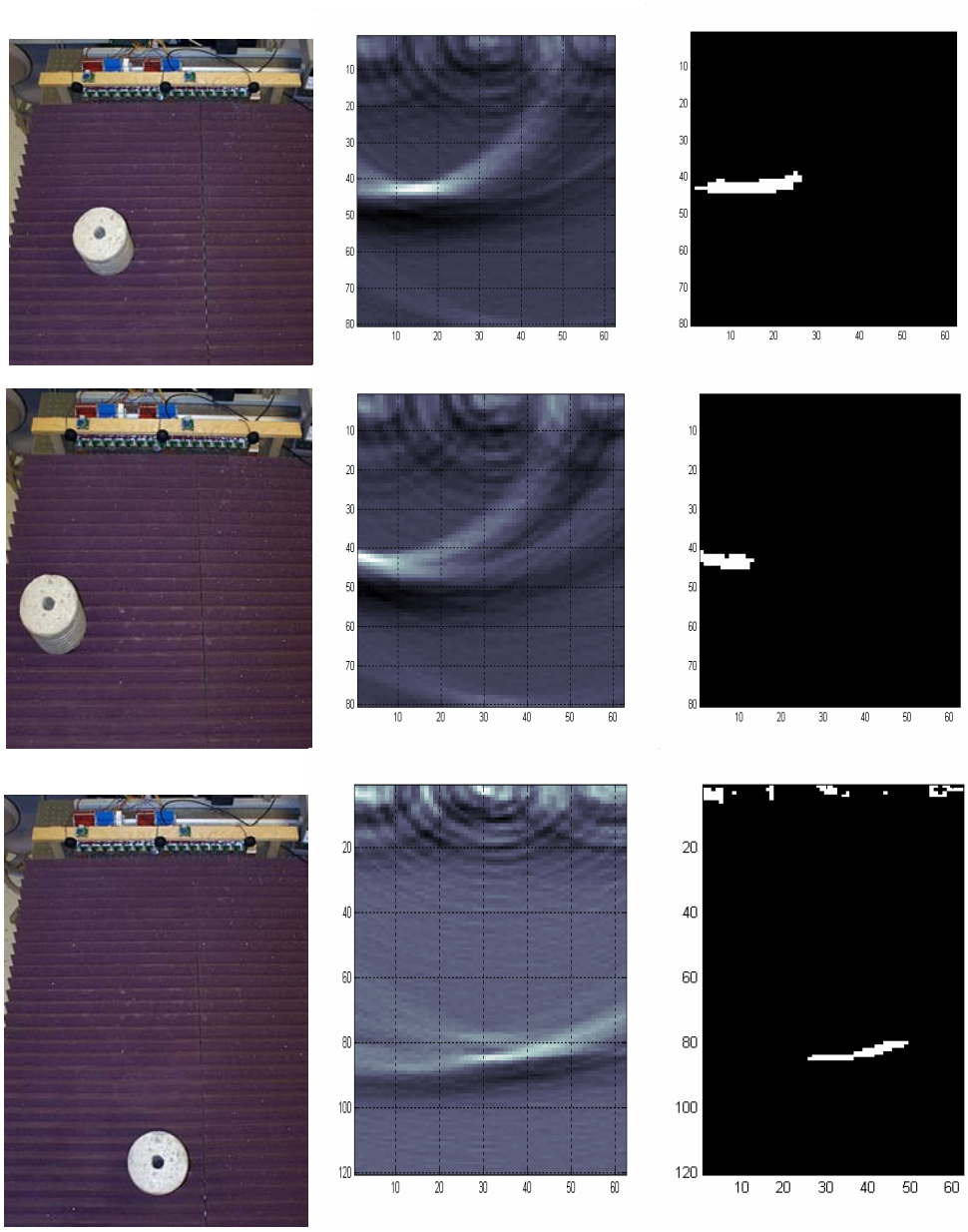


Figure 57. Top view photos and results for concrete cylinder in many positions

The test performed in Figure 57 was designed to support the relationship seen in Figure 43. The number of positions tested, however, was perhaps too small to make any hard conclusions about the accuracy of the resolution inferences made from Figure 43. Just the same, the results of Figure 36 provide some data about the imager performance given target location. For instance, though the third example image (third from top) is partially out of frame, the other three target locations suggest that the resolution does not deviate terribly for different positions of the concrete cylinder. The concrete cylinder, with diameter 15cm, is represented as being between 22cm and 25cm in width for the three target positions shown to be in frame. However, this result may be misleading since the width of the binary image is highly dependent on the threshold value used to generate it. The last setup showing the target at twice the distance did not produce as strong of a raw image; this required the threshold to be reduced by half before production of the binary image. The first three setups in Figure 57 as well as all images in the previous examples (Figures 50-55) used the same threshold whereby each example shows the target to be roughly the same distance from the array. The

results of Figure 57 suggest the need for a gain to be applied to the raw image before thresholding; this gain would be dependent on the distance of the target from the linear array.

Given the positive results shown by indoor testing, the imaging system was mounted to a cart and moved outdoors for outdoor testing. Photos of the cart with linear array and equipment are given in Figure 58 below.

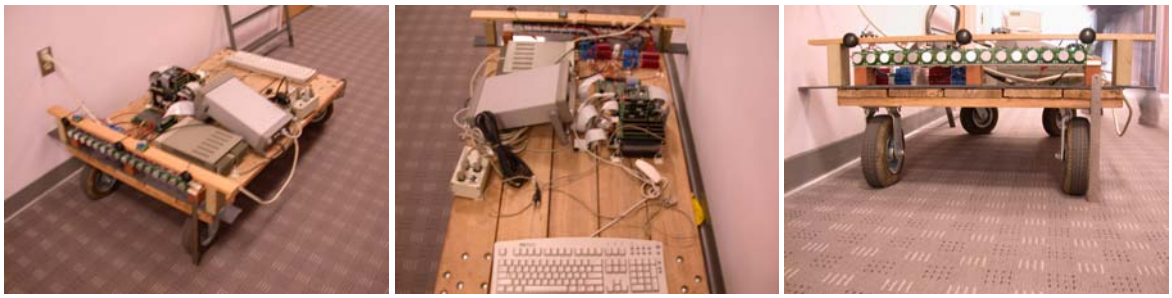
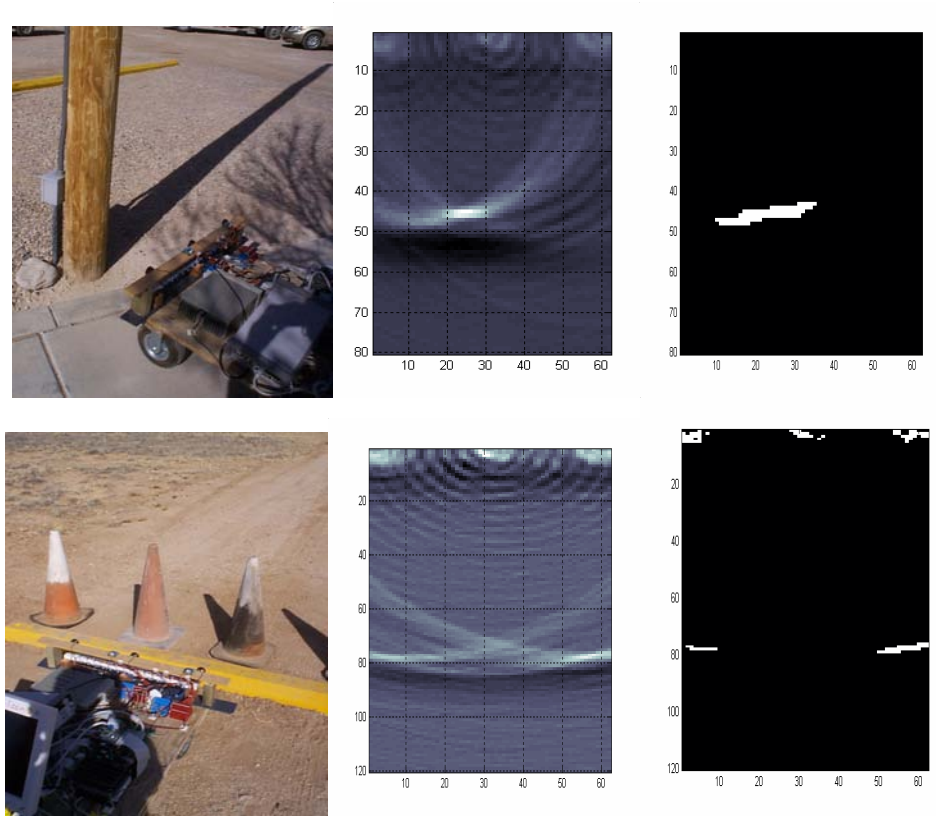


Figure 58. Photos of cart with equipment used for outdoor testing

Unfortunately, outdoor testing was not very successful. After much testing, the experiments failed to generate good images of low-lying objects such as rocks or cactai. It was discovered later that the height of the array above the ground coupled with the narrow beam angle of the individual transducers did not place these low-lying objects into the field of view. As predicted with indoor testing, targets with small diameter media, such as bushes, rebar, or chain-link fences, did not produce good returns. Figure 59 presents some results of outdoor testing.



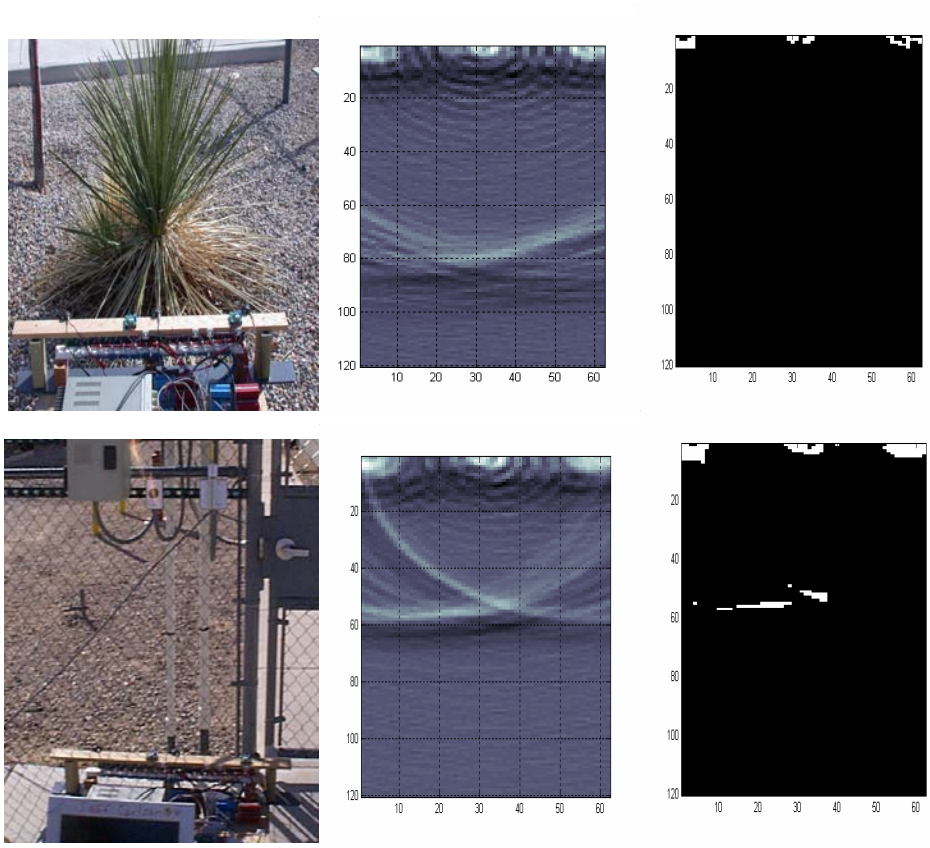


Figure 59. Photos and results for outdoor testing

Obstacle Labeling

To capitalize on the depth and density inference capabilities of the acoustic imager, an algorithm was developed to locate, label, and categorize obstacles that have been identified through the generation of a binary image. The method was then implemented using a custom Matlab routine.

In order to develop a proper algorithm for the discrimination and labeling of obstacles, the image connectedness and stream direction must be established. To simplify this, it was decided that the object field would be 8-connected; each pixel would be connected to each of its 4 edge neighbors and 4 corner neighbors. For this algorithm, there is no background field. The stream direction would be determined based on typical characteristics seen in the acoustic images. Since the object-labeling algorithm will be developed using a sliding window, it is sensitive to upstream facing concavities, and since acoustic images are prone to having these concavities facing the imager axis, it was decided the stream direction would be chosen to be first from side to side, traversing the x-axis positively and then from farthest away to closest, traversing the y-axis negatively . See Figure 60 [16].

With the connectedness and stream direction established, the nature of this sliding window may be explained. The window is a simple four-pixel window with each pixel labeled A, B, C, and D. Most assignments will be made to the A window based on object labels in the neighboring B, C, and D windows.

If the A pixel is 0, no assignments are made, and the window is shifted. If the A pixel equals 1 and B, C, D are all zero, then the A pixel is the beginning of a new object, and it is given a new label. If the A pixel is 1 and either of B, C, D are also 1, then the label attached to A is matched with the label attached to any of the neighboring pixels equaling 1 with the priority B, C, D. For example, if A is 1 and B is also 1, the label

assigned to B will also be assigned to A regardless of the values or labels associated with C and D. Likewise, if A is 1, B is 0, C is 1, and D is 1, the label assigned to C would be copied to A [16].

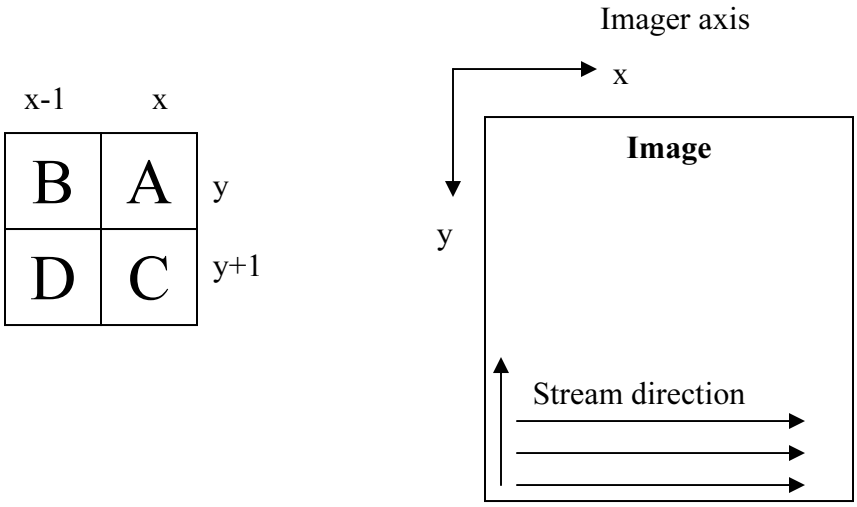
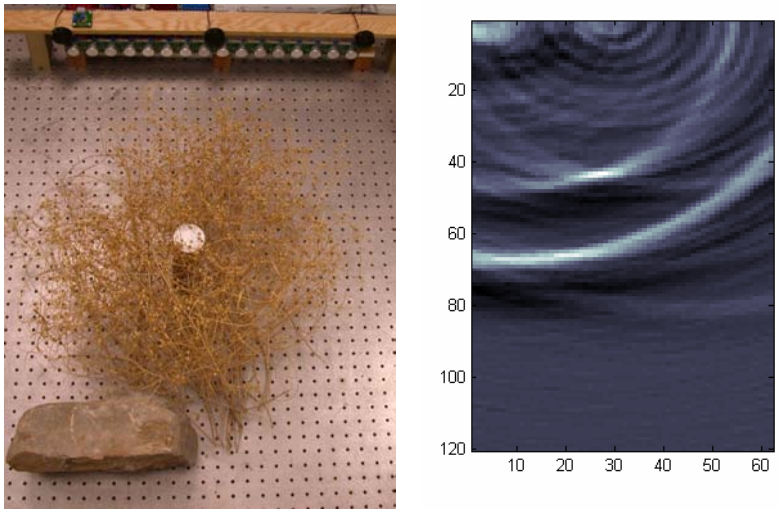


Figure 60. Shows sliding window, image axes, and stream direction

This algorithm can produce some problems, however; it can "split" an object, assigning some pixels to one object label and others to another object label. To reduce this occurrence, a correction routine was added to recursively reassign objects that are connected but separately labeled.

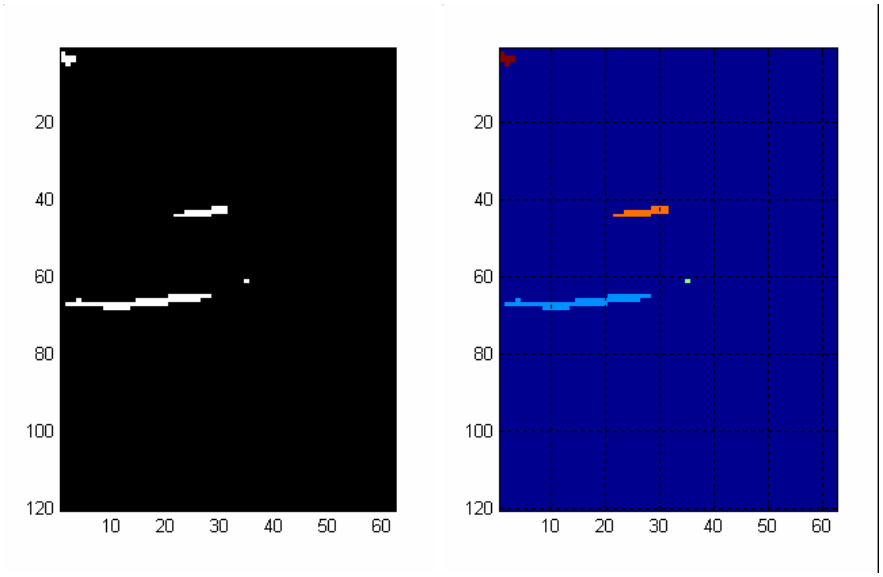
An example of this labeling scheme revisits the rock, tumbleweed, and steel bar obstacle combination shown in the previous section. Figure 61 shows the obstacle photo and the resulting raw image.



(a) (b)

Figure 61. Shows (a) Obstacle photo and (b) Raw acoustic image (not scaled)

Before the algorithm can begin, the raw image is processed into a binary image shown in Figure 62(a). This binary image is then subjected to the object discrimination and labeling algorithm described above. A new image showing the different objects discriminated by color is shown in Figure 62(b).



(a) (b)

Figure 62. Shows (a) the binary image and (b) same image with objects labeled by color

When the algorithm is finished, the Matlab routine generates the following output to describe the objects seen in the acoustic image:

Number_of_Objects =

4

Object	Distance	Bearing	Area	Classification
1.0000	68.3137	-7.8993	45.0000	1.0000
2.0000	70.3278	9.3099	1.0000	3.0000
3.0000	50.9329	2.5758	18.0000	1.0000
4.0000	3.9706	-81.3935	8.0000	2.0000

Classification Table...
Hard Obstacle 1
Sidelobe Effect 2
Ellipse Effect 3

Classifications are assigned to the objects based on their area, distance from origin, and bearing values. For example, the small objects closest to the imager are classified as side lobe noise caused by the receivers being so close to the transmitters. Small objects near large objects are classified as ellipse effects.

The output produced by the algorithm serves to compress the information in the image to a simple data stream that may be used by a computer.

This algorithm was applied to several acoustic images. Its performance was very good generally, however, it had great sensitivity to the threshold used for creating the binary image. If the threshold was too low, the elliptical and sidelobe effects produced a large number of small objects, slowing the program and

generating a lot of information. It was determined that this algorithm could thus be used for threshold calibration. The algorithm could recursively count and classify objects within an image; if the number of objects with Area <3, for instance, was high, the algorithm would reduce the threshold. If the number of such objects were low, the algorithm would increase the threshold. This calibration approach was not employed, but it may be considered for future efforts.

This report surveyed methods of generating acoustic images, some applicable issues and analysis, and the results of hardware testing. In addition, this report discussed image-processing approaches for enhancement and interpretation of acoustic images. Both the acoustic imaging and image processing methods had limited success, but their results support further study. In addition, the results of this effort have generated several ideas as well as background for more focused projects in the future.

References

1. Massa Products Corporation, Massasonic M-5000/Smart Ultrasonic sensor 95 and TR-89/B type 23 transducer http://www.massa.com

2. Simon,D.,Neely,J.2002 "An Investigation into the use of Ultrasonics for Obstacle Characterization in an Unstructured Environment for Autonomous Robotic Navigation" intern symposium paper at Sandia National Labs, Dept. 15212

3. The Mathworks Inc., MATLAB-The language of Technical Computing, Version 6 release 12, September 2000, www.mathworks.com

4. Fink, Kevin. 1994." Computer Simulation of Pressure Fields Generated by Acoustic Lens Beamformers." a thesis at the University of Washington. http://www.fink.com/thesis/title.html

5. Davidson, M, "Sound Reflection" http://freespace.virgin.net/mark.davidson3/reflection/reflection.html

6. Achenbauch,J.D. 1973. Wave Propagation in Elastic Solids. Northwestern University, Evanston,IL. North-Holland Pub.

7. E.A. Ginzel. "The e-Journal of Nondestructive Testing – Ultrasonic Testing" Prometheus Press, Canada. http://www.ndt.net/article/az/ut/append_a/append_a.htm

8. Belcher, E.O., Dinh, H.Q., Lynn, D.C., Laughlin, T.J. "Beamforming and Imaging with Acoustic Lenses in Small High-Frequency Sonars." Applied Physics Laboratory, University of Washington. http://www.apl.washington.edu/programs/DIDSON/Media/acoustic_lenses.pdf

9. "Constant focal length acoustic lens," Patent # 4970705, http://www.anidea.com/Asset/display?asset_id=ANI5050

10. Spletzer,B. 2003. personal communication, Sandia National Labs, Dept. 15212

11. "Acoustic Lens", *Physical Review Letters*, 14 January 2002. http://www.aip.org/mgr/png/2002/146.htm

12. Swanson, D.*C., Signal Processing for Intelligent Sensor Systems*, Marcel Dekker inc., New York, 2000

13. Sanny,J., Moebs,W. *University Physics*, WCB Pub., DuBuque, IA, 1996.

14. Fowler, J.2003. **"**Design and Construction of an Acoustic Imaging Device for Mobile Robotics", viewgraph presentation, Dept 15212, Sandia National Labs

15. Synge,J.L., Schild, A., *Tensor Calculus*, Dover Pub. Inc., New York, 1949

16. Horn, B.C.P., Robot Vision*, McGraw Hill, MIT Press, Cambridge, MA, 1988.*

Distribution

1	MS 0323	D. L. Chavez 01011
1	MS 1002	S. C. Roehrig 15200
1	MS 1003	J. T. Feddema 15211
1	MS 1003	D. D. Padilla 15211
1	MS 1003	J. J. Carlson 15211
1	MS 1003	J. Neely 15211
1	MS 1003	D. K. Novick 15211
1	MS 1003	P. A. Davidson Jr. 15211
6	MS 1003	B. Ewing 15211
1	MS 9018	Central Technical Files, 8945-1
2	MS 0899	Technical Library, 9616